

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
31 March 2005 (31.03.2005)

PCT

(10) International Publication Number  
**WO 2005/028625 A2**

(51) International Patent Classification<sup>7</sup>: **C12N**  
  
(21) International Application Number:  
PCT/US2004/030397

(US). **GAO, Feng** [US/US]; 702 Country Club Drive,  
Durham, NC 27712 (US). **LIAO, Hua-Xin** [CN/US]; 200  
Pebble Springs Road, Chapel Hill, NC 27514 (US).

(22) International Filing Date:  
17 September 2004 (17.09.2004)

(74) Agent: **WILSON, Mary, J.**; Nixon & Vanderhye P.C.,  
1100 North Glebe Road, Suite 800, Arlington, VA 22201-  
4714 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/503,460 17 September 2003 (17.09.2003) US  
60/604,722 27 August 2004 (27.08.2004) US

(71) Applicant (for all designated States except US): **DUKE UNIVERSITY** [US/US]; P.O. Box 90083, Durham, NC 27708-0083 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(71) Applicants and

(72) Inventors: **KORBER, Bette, T.** [US/US]; 1290 Big Rock Loop, Los Alamos, NM 87544 (US). **HAHN, Beatrice, H.** [DE/US]; 3571 Rockhill Road, Birmingham, AL 35223 (US). **SHAW, George, M.** [US/US]; 3571 Rockhill Road, Birmingham, AL 35223 (US). **KOTHE, Denise** [US/US]; 734 39th Street, Birmingham, AL 35222 (US). **LI, Ying Ying** [CN/US]; 2529 Mountain Cove, Hoover, AL 35266 (US). **DECKER, Julie** [US/US]; 1116 Colonial Drive, Alabaster, AL 35007 (US).

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Published:

— without international search report and to be republished upon receipt of that report

(72) Inventors; and

(75) Inventors/Applicants (for US only): **HAYNES, Barton, F.** [US/US]; 3923 Wentworth Street, Durham, NC 27707

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **CONSENSUS/ANCESTRAL IMMUNOGENS**

(57) Abstract: The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralizes a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.



WO 2005/028625 A2

THIS PAGE BLANK (USPTO)

## CONSENSUS/ANCESTRAL IMMUNOGENS

This application claims priority from Prov. Appln. No. 60/503,460, filed September 17, 2003, and Prov. Appln. No. 60/604,722, filed August 27, 2004, the entire contents of which are incorporated herein by reference.

### TECHNICAL FIELD

The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.

### BACKGROUND

The high level of genetic variability of HIV-1 has presented a major hurdle for AIDS vaccine development. Genetic differences among HIV-1 groups M, N, and O are extensive, ranging from 30% to 50% in gag and env genes, respectively (Gurtler et al, J. Virol. 68:1581-1585 (1994), Vanden Haesevelde et al, J. Virol. 68:1586-1596 (1994), Simon et al, Nat. Med. 4:1032-1037 (1998), Kuiken et al, Human

retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid sequences (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico)). Viruses within group M are further classified into nine genetically distinct subtypes (A-D, F-H, J and K) (Kuiken et al, Human retroviruses and AIDS 2000: a compilation and analysis of nucleic acid and amino acid sequences (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, Robertson et al, Science 288:55-56 (2000), Robertson et al, Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (2000)). With the genetic variation as high as 30% in *env* genes among HIV-1 subtypes, it has been difficult to consistently elicit cross-subtype T and B cell immune responses against all HIV-1 subtypes. HIV-1 also frequently recombines among different subtypes to create circulating recombinant forms (CRFs) (Robertson et al, Science 288:55-56 (2000), Robertson et al, Human retroviruses and AIDS 1999: a compilation and analysis of nucleic acid and amino acid sequences, eds. Kuiken et al (Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico), pp. 492-505 (2000), Carr et al, Human retroviruses and AIDS 1998: a compilation and analysis of nucleic acid and



amino acid sequences, eds. Korber et al (Theoretical  
Biology and Biophysics Group, Los Alamos National  
Laboratory, Los Alamos, New Mexico), pp. III-10-III-  
19 (1998)). Over 20% of HIV-1 isolates are  
5 recombinant in geographic areas where multiple  
subtypes are common (Robertson et al, Nature  
374:124-126 (1995), Cornelissen et al, J. virol.  
70:8209-8212 (1996), Dowling et al, AIDS 16:1809-  
1820 (2002)), and high prevalence rates of  
10 recombinant viruses may further complicate the  
design of experimental HIV-1 immunogens.

To overcome these challenges in AIDS vaccine  
development, three computer models (consensus,  
ancestor and center of the tree) have been used to  
15 generate centralized HIV-1 genes to (Gaschen et al,  
Science 296:2354-2360 (2002), Gao et al, Science  
299:1517-1518 (2003), Nickle et al, Science  
299:1515-1517 (2003), Novitsky et al, J. Virol.  
76:5435-5451 (2002), Ellenberger et al, Virology  
20 302:155-163 (2002), Korber et al, Science 288:1789-  
1796 (2000)). The biology of HIV gives rise to  
star-like phylogenies, and as a consequence of this,  
the three kinds of sequences differ from each other  
by 2 - 5% (Gao et al, Science 299:1517-1518 (2003)).  
25 Any of the three centralized gene strategies will  
reduce the protein distances between immunogens and  
field virus strains. Consensus sequences minimize  
the degree of sequence dissimilarity between a  
vaccine strain and contemporary circulating viruses  
30 by creating artificial sequences based on the most  
common amino acid in each position in an alignment

(Gaschen et al, Science 296:2354-2360 (2002)).  
Ancestral sequences are similar to consensus  
sequences but are generated using maximum-likelihood  
phylogenetic analysis methods (Gaschen et al,  
5 Science 296:2354-2360 (2002), Nickle et al, Science  
299:1515-1517 (2003)) . In doing so, this method  
recreates the hypothetical ancestral genes of the  
analyzed current wild-type sequences (Figure 26).  
Nickle et al proposed another method to generate  
10 centralized HIV-1 sequences, center of the tree  
(COT), that is similar to ancestral sequences but  
less influenced by outliers (Science 299:1515-1517  
(2003)).

The present invention results, at least in  
15 part, from the results of studies designed to  
determine if centralized immunogens can induce both  
T and B cell immune responses in animals. These  
studies involved the generation of an artificial  
group M consensus env gene (CON6), and construction  
20 of DNA plasmids and recombinant vaccinia viruses to  
express CON6 envelopes as soluble gp120 and gp140CF  
proteins. The results demonstrate that CON6 Env  
proteins are biologically functional, possess  
linear, conformational and glycan-dependent epitopes  
25 of wild-type HIV-1, and induce cytokine-producing T  
cells that recognize T cell epitopes of both HIV  
subtypes B and C. Importantly, CON6 gp120 and  
gp140CF proteins induce antibodies that neutralize  
subsets of subtype B and C HIV-1 primary isolates.

30 The iterative nature of study of the  
centralized HIV-1 gene approach is derived from the

rapidly expanding evolution of HIV-1 sequences, and the fact that sequences collected in the HIV sequence database (that is, the Los Alamos National Database) are continually being updated with new sequences each year. The CON6 gp120 envelope gene derives from Year 1999 Los Alamos National Database sequences, and Con-S derives from Year 2000 Los Alamos National Database sequences. In addition, CON6 has Chinese subtype C V1, V2, V4, and V5 Env sequences, while Con-S has all group M consensus Env constant and variable regions, that have been shortened to minimal-length variable loops. Codon-optimized genes for a series of Year 2003 group M and subtype consensus sequences have been designed, as have a corresponding series of wild-type HIV-1 Env genes for comparison, for use in inducing broadly reactive T and B cell responses to HIV-1 primary isolates.

#### SUMMARY OF THE INVENTION

The present invention relates to an immunogen for inducing antibodies that neutralize a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response, and to nucleic acid sequences encoding same. The invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using such an immunogen.

Objects and advantages of the present invention will be clear from the description that follows.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1D: Generation and expression of the group M consensus env gene (CON6). The complete amino acid sequence of CON6 gp160 is shown.

5 (Fig. 1A) The five regions from the wild-type CRF08\_BC (98CN006) env gene are indicated by underlined letters. Variable regions are indicated by brackets above the sequences. Potential N-linked glycosylation sites are highlighted with bold-faced  
10 letters. (Fig. 1B) Constructs of CON6 gp120 and gp140CF. CON6 gp120 and gp140CF plasmids were engineered by introducing a stop codon after the gp120 cleavage site or before the transmembrane domain, respectively. The gp120/gp41 cleavage site  
15 and fusion domain of gp41 were deleted in the gp140CF protein. (Fig. 1C) Expression of CON6 gp120 and gp140CF. CON6 gp120 and gp140CF were purified from the cell culture supernatants of rVV-infected 293T cells with *galanthus Nivalis argarose* lectin  
20 columns. Both gp120 and gp140CF were separated on a 10% SDS-polyarylamide gel and stained with Commassie blue. (Fig. 1D.) CON6 env gene optimized based on codon usage for highly expressed human genes.

Figures 2A-2E. Binding of CON6 gp120 gp140 CF  
25 to soluble CD4 (sCD4) and anti-Env mAbs. (Figs. 2A-2B) Each of the indicated mabs and sCD4 was covalently immobilized to a CM5 sensor chip (BIAcore) and CON6 gp120 (Fig. 2A) or gp140CF (Fig.

2B) (100  $\mu\text{g/ml}$  and 300  $\mu\text{g/ml}$ , respectively) were injected over each surface. Both gp120 and gp140CF proteins reacted with each anti-gp120 mabs tested except for 17b mab, which showed negligible binding to both CON6 gp120 and gp140CF. To determine induction of 17b mab binding to CON6 gp120 and gp140CF, CON6 gp120 (Fig. 2C) or gp140CF (Fig. 2D) proteins were captured (400-580 RU) on individual flow cells immobilized with sCD4 or mabs A32 or T8. Following stabilization of each of the surface, mAb 17b was injected and flowed over each of the immobilized flow cells. Overlay of curves show that the binding of mab 17b to CON6 Env proteins was markedly enhanced on both sCD4 and mab A32 surfaces but not on the T8 surface (Figs. 2C-2D). To determine binding of CON6 gp120 and gp140CF to human mabs in ELISA, stock solutions of 20 $\mu\text{g/ml}$  of mabs 447, F39F, A32, IgG1b12 and 2F5 on CON6 gp120 and gp140CF were tittered (Fig. 2E). Mabs 447 (V3), F39F (V3) A32 (gp120) and IgG1b12 (CD4 binding site) each bound to both CON6 gp120 and 140 well, while 2F5 (anti-gp41 ELDKWS) only bound gp140CF. The concentration at endpoint titer on gp120 for mab 447 and F39F binding was <0.003  $\mu\text{g/ml}$  and 0.006  $\mu\text{g/ml}$ , respectively; for mab A32 was <0.125  $\mu\text{g/ml}$ ; for IgG1b12 was <0.002  $\mu\text{g/ml}$ ; and for 2F5 was 0.016  $\mu\text{g/ml}$ .

Figures 3A and 3B. Infectivity and coreceptor usage of CON6 envelope. (Fig. 3A) CON6 and control

env plasmids were cotransfected with HIV-1/SG3Δenv backbone into human 293T cells to generate Env-pseudovirions. Equal amounts of each pseudovirion (5 ng p24) were used to infect JC53-BL cells. The infectivity was determined by counting the number of blue cells (infectious units, IU) per microgram of p24 of pseudovirions (IU/μg p24) after staining the infected cells for β-gal expression. (Fig. 3B) Coreceptor usage of the CON6 env gene was determined on JC53BL cells treated with AMD3100 and/or TAK-799 for 1 hr (37°C) then infected with equal amounts of p24 (5 ng) of each Env-pseudovirion. Infectivity in the control group (no blocking agent) was set as 100%. Blocking efficiency was expressed as the percentage of IU from blocking experiments compared to those from control cultures without blocking agents. Data shown are mean ± SD.

Figure 4. Western blot analysis of multiple subtype Env proteins against multiple subtype antisera. Equal amount of Env proteins (100 ng) were separated on 10% SDS-polyacrylamide gels. Following electrophoresis, proteins were transferred to Hybond ECL nitrocellulose membranes and reacted with sera from HIV-1 infected patients (1:1,000) or guinea pigs immunized with CON6 gp120 DNA prime, rVV boost (1:1,000). Protein-bound antibody was probed with fluorescent-labeled secondary antibodies and the images scanned and recorded on an infrared imager Odyssey (Li-Cor, Lincoln, NE). Subtypes are

indicated by single-letters after Env protein and serum IDs. Four to six sera were tested for each subtype, and reaction patterns were similar among all sera from the same subtype. One representative  
5 result for each subtype serum is shown.

Figure 5. T cell immune responses induced by CON6 Env immunogens in mice. Splenocytes were isolated from individual immunized mice (5 mice/group). After splenocytes were stimulated in  
10 vitro with overlapping Env peptide pools of CON6 (black column), subtype B (hatched column), subtype C (white column), and medium (no peptide; gray column), INF- $\gamma$  producing cells were determined by the ELISPOT assay. T cell IFN- $\gamma$  responses induced  
15 by either CON6 gp120 or gp140CF were compared to those induced by subtype specific Env immunogens (JRFL and 96ZM651). Total responses for each envelope peptide pool are expressed as SFCs per million splenocytes. The values for each column are  
20 the mean  $\pm$  SEM (of IFN- $\gamma$  SFCs (n=5 mice/group)).

Figures 6A-6E. Construction of codon usage optimized subtype C ancestral and consensus envelope genes (Figs. 6A and 6B, respectively). Ancestral and consensus amino acid sequences (Figs. 6C and 6D,  
25 respectively) were transcribed to mirror the codon usage of highly expressed human genes. Paired oligonucleotides (80-mers) overlapping by 20 bp were designed to contain 5' invariant sequences including

the restriction enzyme sites EcoRI, BbsI, Bam HI and BsmBI. BbsI and BsmBI are Type II restriction enzymes that cleave outside of their recognition sequences. Paired oligomers were linked  
5 individually using PCR and primers complimentary to the 18 bp invariant sequences in a stepwise fashion, yielding 140bp PCR products. These were subcloned into pGEM-T and sequenced to confirm the absence of inadvertant mutations/deletions. Four individual  
10 pGEM-T subclones containing the proper inserts were digested and ligated together into pcDNA3.1. Multi-fragment ligations occurred repeatedly amongst groups of fragments in a stepwise manner from the 5' to the 3' end of the gene until the entire gene was  
15 reconstructed in pcDNA3.1. (See schematic in Fig. 6E.)

Figure 7. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also contain  
20 the reporter cassettes of luciferase and  $\beta$ -galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24 or 96-well plates, incubated at 37°C  
25 for 24 hours and treated with DEAE-Dextran at 37°C for 30 minutes. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-Dextran, and allowed to incubate for 3 hours at 37°C after which an additional cell media is added to each



well. Following a final 48-hour incubation at 37°C, cells are either fixed, stained using X-Gal to visualize  $\beta$ -galactosidase expressing blue foci or frozen-thawed three times to measure luciferase activity.

Figure 8. Sequence alignment of subtype C ancestral and consensus *env* genes. Alignment of the subtype C ancestral (bottom line) and consensus (top line) *env* sequences showing a 95.5% sequence homology; amino acid sequence differences are indicated. One noted difference is the addition of a glycosylation site in the C ancestral *env* gene at the base of the V1 loop. A plus sign indicates a within-class difference of amino acid at the indicated position; a bar indicates a change in the class of amino acid. Potential N-glycosylation sites are marked in blue. The position of truncation for the *gp140* gene is also shown.

Figure 9. Expression of subtype C ancestral and consensus envelopes in 293T cells. Plasmids containing codon-optimized *gp160*, *gp140*, or *gp120* subtype C ancestral and consensus genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2  $\mu$ g of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins

were transferred to a PVDF membrane and probed with HIV-1 plasma from a subtype C infected patient.

Figures 10A and 10B. Fig. 10A. *Trans* complementation of *env*-deficient HIV-1 with codon-optimized subtype C ancestral and consensus gp160 and gp140. Plasmids containing codon-optimized, subtype C ancestral or consensus *gp160* or *gp140* genes were co-transfected into 293T cells with an HIV-1/SG3Δ*env* provirus. 48 hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified by centrifugation, filtered through a 0.2μM filter, and pelleted through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel for particles containing a codon-optimized envelope. 250ng of p24 was loaded per lane for particles generated by co-transfection of a rev-dependent wild-type subtype C 96ZAM651*env* gene. Differences in the amount of p24 loaded per lane were necessary to ensure visualization of the rev-dependent envelopes by Western Blot. Proteins were transferred to a PVDF membrane and probed with pooled plasma from HIV-1 subtype B and subtype C infected individuals. Fig. 10B. Infectivity of virus particles containing subtype C ancestral and consensus envelope glycoproteins. Infectivity of pseudotyped virus containing ancestral or consensus *gp160* or *gp140* envelope was determined using the

JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize  $\beta$ -galactosidase expressing cells. Infectivity is represented as infectious units per ng of p24 to normalize for differences in the concentration of the input pseudovirions.

Figure 11. Co-receptor usage of subtype C ancestral and consensus envelopes. Pseudotyped particles containing ancestral or consensus envelope were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), or AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4, and YU-2, a known CCR5-using isolate, were included as controls.

Figures 12A-12C. Neutralization sensitivity of subtype C ancestral and consensus envelope glycoproteins. Equivalent amounts of pseudovirions containing the ancestral, consensus or 96ZAM651 *gp160* envelopes (1,500 infectious units) were pre-incubated with a panel of plasma samples from HIV-1 subtype C infected patients and then added to the JC53-BL cell monolayer in 96-well plates. Plates

were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity is calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC<sub>50</sub>) and the actual % neutralization at each antibody dilution are then calculated for each virus. The results of all luciferase experiments are confirmed by direct counting of blue foci in parallel infections.

Figures 13A-13F. Protein expression of consensus subtype C Gag (Fig. 13A) and Nef (Fig. 13B) following transfection into 293T cells. Consensus subtype C Gag and Nef amino acid sequences are set forth in Figs. 13C and 13D, respectively, and encoding sequences are set forth in Figs. 13E and 13F, respectively.

Figures 14A-14C. Figs. 14A and 14B show the Con-S Env amino acid sequence and encoding sequence, respectively. Fig. 14C shows expression of Group M consensus Con-S Env proteins using an *in vitro* transcription and translation system.

Figures 15A and 15B. Expression of Con-S env gene in mammalian cells. (Fig. 15A - cell lysate, Fig. 15B - supernatant.)

Figures 16A and 16B. Infectivity (Fig. 16A) and coreceptor usage (Fig. 16B) of CON6 and Con-S env genes.

Figures 17A-17C. Env protein incorporation in  
5 CON6 and Con-S Env-pseudovirions. (Fig. 17A - lysate, Fig. 17B - supernatant, Fig. 17C pellet.)

Figures 18A-18D. Figs. 18A and 18B show subtype A consensus Env amino acid sequence and nucleic acid sequence encoding same, respectively.  
10 Figs. 18C and 18D show expression of A.con env gene in mammalian cells (Fig. 18C - cell lysate, Fig. 18D - supernatant).

Figures 19A-19H. M.con.gag (Fig. 19A), M.con.pol (Fig. 19B), M.con.nef (Fig. 19C) and  
15 C.con.pol (Fig. 19D) nucleic acid sequences and corresponding encoded amino acid sequences (Figs. 19E-19H, respectively).

Figures 20A-20D. Subtype B consensus gag (Fig. 20A) and env (Fig. 20B) genes. Corresponding amino  
20 acid sequences are shown in Figs. 20C and 20D.

Figure 21. Expression of subtype B consensus env and gag genes in 293T cells. Plasmids containing codon-optimized subtype B consensus gp160, gp140, and gag genes were transfected into  
25 293T cells, and protein expression was examined by

Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2  $\mu$ g of total protein was loaded per lane  
5 on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.

Figure 22. Co-receptor usage of subtype B consensus envelopes. Pseudotyped particles  
10 containing the subtype B consensus gp160 Env were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-receptor usage.  
15 NL4.3, an isolate known to utilize CXCR4 and YU-2, a known CCR5-using isolate, were included as controls.

Figures 23A and 23B. Trans complementation of env-deficient HIV-1 with codon-optimized subtype B consensus gp160 and gp140 genes. Plasmids  
20 containing codon-optimized, subtype B consensus gp160 or gp140 genes were co-transfected into 293T cells with an HIV-1/SG3 $\Delta$ env provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a  
25 tabletop centrifuge, filtered through a 0.2 $\mu$ M filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen

assay; 25 ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. *Trans*  
5 complementation with a rev-dependent NL4.3 env was included for control. Figure 23B. Infectivity of virus particles containing the subtype B consensus envelope. Infectivity of pseudotyped virus containing consensus B gp160 or gp140 was determined  
10 using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation  
15 period, cells were fixed and stained to visualize  $\beta$ -galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

Figures 24A-24D. Neutralization sensitivity of virions containing subtype B consensus gp160  
20 envelope. Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1  
25 subtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by

dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration ( $IC_{50}$ ) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections. Fig. 24A. Neutralization of Pseudovirions containing Subtype B consensus Env (gp160). Fig. 24B. Neutralization of Pseudovirions containing NL4.3 Env (gp160). Fig. 24C. Neutralization of Pseudovirions containing Subtype B consensus Env (gp160). Fig. 24D. Neutralization of Pseudovirions containing NL4.3 Env (gp160).

Figures 25A and 25B. Fig. 25A. Density and p24 analysis of sucrose gradient fractions. 0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml. Fig. 25B. VLP production by co-transfection of subtype B consensus gag and env genes. 293T cells were co-transfected with subtype B consensus gag and



env genes. Cell supernatants were harvested 48-  
hours post-transfection, clarified through at 20%  
sucrose cushion, and further purified through a 20-  
60% sucrose gradient. Select fractions from the  
5 gradient were pooled, added to 20ml of PBS, and  
centrifuged overnight at 100,000 x g. Resuspended  
pellets were loaded onto a 4-20% SDS-PAGE gel,  
proteins were transferred to a PVDF membrane, and  
probed with plasma from an HIV-1 subtype B infected  
10 individual.

Figures 26A and 26B. Fig. 26A. 2000 Con-S  
140CFI.ENV. Fig. 26B. Codon-optimized Year 2000  
Con-S 140CFI.seq.

Figure 27. Individual C57BL/6 mouse T cell  
15 responses to HIV-1 envelope peptides. Comparative  
immunogenicity of CON6 gp140CFI and Con-S gp140CFI  
in C57BL/C mice. Mice were immunized with either  
HIV5305 (Subtype A), 2801 (Subtype B), CON6 or Con-S  
Envelope genes in DNA prime, rVV boost regimens, 5  
20 mice per group. Spleen cells were assayed for IFN- $\gamma$   
spot-forming cells 10 days after rVV boost, using  
mixtures of overlapping peptides from Envs of HIV-1  
UG37(A), MN(B), Ch19(C), 89.6(B) SF162(B) or no  
peptide negative control.

25 Figures 28A-28C. Fig. 28A. Con-B 2003 Env. pep  
(841 a.a.). Amino acid sequence underlined is the  
fusion domain that is deleted in 140CF design and

the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 28B. Con-B-140CF.pep (632 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 28C. Codon-optimized Con-B 140CF.seq (1927 nt.).

Figures 29A-29C. Fig. 29A. CON\_OF\_CONS-2003 (829 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 29B. Cons-2003 140CF.pep (620 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 29C. CODON-OPTIMIZED Cons-2003 140CF.seq (1891 nt.).

Figures 30A-30C. Fig. 30A. CONSENSUS\_A1-2003 (845 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 30B. Con-A1-2003 140CF.pep (629 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 30C. CODON-OPTIMIZED Con-A1-2003.seq.

Figures 31A-31C. Fig. 31A. CONSENSUS\_C-2003 (835 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 31B. Con-C 2003 140CF.pep (619 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 31C. CODON-OPTIMIZED Con-C-2003 (140 CF (1,888 nt.)).

Figures 32A-32C. Fig. 32A. CONSENSUS\_G-2003 (842 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 32B. Con-G-2003 140CF.pep (626 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage site. Fig. 32C. CODON-OPTIMIZED Con-G-2003.seq.

Figures 33A-33C. Fig. 33A. CONSENSUS\_01\_AE-2003 (854 a.a.). Amino acid sequence underlined is the fusion domain that is deleted in 140CF design and the "W" underlined is the last amino acid at the C-terminus, all amino acids after the "W" are deleted in the 140CF design. Fig. 33B. Con-AE01-2003 140CF.pep (638 a.a.). Amino acids in bold identify the junction of the deleted fusion cleavage

site. Fig. 33C. CODON-OPTIMIZED Con-AE01-2003.seq.  
(1945 nt.).

Figures 34A-34C. Fig. 34A. Wild-type subtype  
A Env. 00KE\_MSA4076-A (Subtype A, 891 a.a.). Amino  
5 acid sequence underlined is the fusion domain that  
is deleted in 140CF design and the "W" underlined  
is the last amino acid at the C-terminus, all amino  
acids after the "W" are deleted in the 140CF design.  
Fig. 34B. 00KE\_MSA4076-A 140CF.pep (647 a.a.).  
10 Amino acids in bold identify the junction of the  
deleted fusion cleavage site. Fig. 34C. CODON-  
OPTIMIZED 00KE\_MSA4076-A 140CF.seq. (1972 nt.).

Figures 35A-35C. Fig. 35A. Wild-type subtype  
B. QH0515.1g gp160 (861 a.a.). Amino acid sequence  
15 underlined is the fusion domain that is deleted in  
140CF design and the "W" underlined is the last  
amino acid at the C-terminus, all amino acids after  
the "W" are deleted in the 140CF design. Fig. 35B.  
QH0515.1g 140CF (651 a.a.). Amino acids in bold  
20 identify the junction of the deleted fusion cleavage  
site. Fig. 35C. CODON-OPTIMIZED QH0515.1g  
140CF.seq (1984 nt.).

Figures 36A-36C. Fig. 36A. Wild-type subtype  
C. DU123.6 gp160 (854 a.a.). Amino acid sequence  
25 underlined is the fusion domain that is deleted in  
140CF design and the "W" underlined is the last  
amino acid at the C-terminus, all amino acids after

the "W" are deleted in the 140CF design. Fig. 36B.  
DU123.6 140CF (638 a.a.). Amino acids in bold  
identify the junction of the deleted fusion cleavage  
site. Fig. 36C. CODON-OPTIMIZED DU123.6 140CF.seq  
5 (1945 nt.).

Figures 37A-37C. Fig. 37A. Wild-type subtype  
CRF01\_AE. 97CNGX2F-AE (854 a.a.). Amino acid  
sequence underlined is the fusion domain that is  
deleted in 140CF design and the "W" underlined is  
10 the last amino acid at the C-terminus, all amino  
acids after the "W" are deleted in the 140CF design.  
Fig. 37B. 97CNGX2F-AE 140CF.pep (629 a.a.). Amino  
acids in bold identify the junction of the deleted  
fusion cleavage site. Fig. 37C. CODON-OPTIMIZED  
15 97CNGX2F-AE 140CF.seq (1921 nt.).

Figures 38A-38C. Fig. 38A. Wild-type DRCBL-G  
(854 a.a.). Amino acid sequence underlined is the  
fusion domain that is deleted in 140CF design and  
the "W" underlined is the last amino acid at the  
20 C-terminus, all amino acids after the "W" are  
deleted in the 140CF design. Fig. 38B. DRCBL-G  
140CF.pep (630 a.a.). Amino acids in bold identify  
the junction of the deleted fusion cleavage site.  
Fig. 38C. CODON-OPTIMIZED DRCBL-G 140CF.seq (1921  
25 nt.).

Figures 39A and 39B. Fig. 39A. 2003 Con-S  
Env. Fig. 39B. 2003 Con-S Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 40A and 40B. Fig. 40A. 2003 M.  
5 Group.Anc Env. Fig. 40B. 2003 M. Group.anc  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

Figures 41A and 41B. Fig. 41A. 2003 CON\_A1  
Env. Fig. 41B. 2003 CON\_A1 Env.seq.opt.  
10 (Seq.opt. = codon optimized encoding sequence.)

Figures 42A and 42B. Fig. 42A. 2003 A1.Anc  
Env. Figs. 42B. 2003 A1.anc Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 43A and 43B. Fig. 43A. 2003 CON\_A2  
15 Env. Fig. 43B. 2003 CON\_A2 Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 44A and 44B. Fig. 44A. 2003 CON\_B  
Env. Fig. 44B. 2003 CON\_B Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 45A and 45B. Fig. 45A. 2003 B.anc  
20 Env. Figs. 45B. 2003 B.anc Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 46A and 46B. Fig. 46A. 2003 CON\_C  
Env. Fig. 46B. 2003 CON\_C Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 47A and 47B. Fig. 47A. 2003 C.anc  
5 Env. Fig. 47B. 2003 C.anc Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 48A and 48B. Fig. 48A. 2003 CON\_D  
Env. Fig. 48B. 2003 CON\_D Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

10 Figures 49A and 49B. Fig. 49A. 2003 CON\_F1  
Env. Fig. 49B. 2003 CON\_F1 Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 50A and 50B. Fig. 50A. 2003 CON\_F2  
Env. Fig. 50B. 2003 CON\_F2 Env.seq.opt.  
15 (Seq.opt. = codon optimized encoding sequence.)

Figures 51A and 51B. Fig. 51A. 2003 CON\_G  
Env. Fig. 51B. 2003 CON\_G Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 52A and 52B. Fig. 52A. 2003 CON\_H  
20 Env. Fig. 52B. 2003 CON\_H Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 53A and 53B. Fig. 53A. 2003 CON\_01\_AE  
Env. Fig. 53B. 2003 CON\_01\_AE Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 54A and 54B. Fig. 54A. 2003 CON\_02\_AG  
5 Env. Fig. 54B. 2003 CON\_02\_AG Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 55A and 55B. Fig. 55A. 2003 CON\_03\_AB  
Env. Fig. 55B. 2003 CON\_03\_AB Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

10 Figures 56A and 56B. Fig. 56A. 2003  
CON\_04\_CPX Env. Fig. 56B. 2003 CON\_04\_CPX  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

Figures 57A and 57B. Fig. 57A. 2003  
15 CON\_06\_CPX Env. Fig. 57B. 2003 CON\_06\_CPX  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

Figures 58A and 58B. Fig. 58A. 2003 CON\_08\_BC  
Env. Fig. 58B. 2003 CON\_08\_BC Env.seq.opt.  
20 (Seq.opt. = codon optimized encoding sequence.)

Figures 59A and 59B. Fig. 59A. 2003 CON\_10\_CD  
Env. Fig. 59B. 2003 CON\_10\_CD Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)



Figures 60A and 60B. Fig. 60A. 2003  
CON\_11\_CPX Env. Fig. 60B. 2003 CON\_11\_CPX  
Env.seq.opt. (Seq.opt. = codon optimized encoding  
sequence.)

5 Figures 61A and 61B. Fig. 61A. 2003 CON\_12\_BF  
Env. Fig. 61B. 2003 CON\_12\_BF Env.seq.opt.  
(Seq.opt. = codon optimized encoding sequence.)

Figures 62A and 62B. Fig. 62A. 2003 CON\_14\_BG  
Env. Fig. 62B. 2003 CON\_14\_BG Env.seq.opt.  
10 (Seq.opt. = codon optimized encoding sequence.)

Figures 63A and 63B. Fig. 63A. 2003\_CON\_S  
gag.PEP. Fig. 63B. 2003\_CON\_S gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 64A and 64B. Fig. 64A.  
15 2003\_M.GROUP.anc gag.PEP. Fig. 64B.  
2003\_M.GROUP.anc gag.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 65A-65D. Fig. 65A. 2003\_CON\_A1  
gag.PEP. Fig. 65B. 2003\_CON\_A1 gag.OPT. Fig. 65C.  
20 2003\_A1.anc gag.PEP. Fig. 65D. 2003\_A1.anc  
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 66A and 66B. Fig. 66A. 2003\_CON\_A2  
gag.PEP. Fig. 66B. 2003\_CON\_A2 gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 67A-67D. Fig. 67A. 2003\_CON\_B  
5 gag.PEP. Fig. 67B. 2003\_CON\_B gag.OPT. Fig. 67C.  
2003\_B.anc gag.PEP. Fig. 67D. 2003\_B.anc gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 68A-68D. Fig. 68A. 2003\_CON\_C  
gag.PEP. Fig. 68B. 2003\_CON\_C gag.OPT. Fig. 68C.  
10 2003\_C.anc.gag.PEP. Fig. 68D. 2003\_C.anc.gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 69A and 69B. Fig. 69A. 2003\_CON\_D  
gag.PEP. Fig. 69B. 2003\_CON\_D gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 70A and 70B. Fig. 70A. 2003\_CON\_F  
15 gag.PEP. Fig. 70B. 2003\_CON\_F gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 71A and 71B. Fig. 71A. 2003\_CON\_G  
gag.PEP. Fig. 71B. 2003\_CON\_G gag.OPT.  
20 (OPT = codon optimized encoding sequence.)

Figures 72A and 72B. Fig. 72A. 2003\_CON\_H  
gag.PEP. Fig. 72B. 2003\_CON\_H gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 73A and 73B. Fig. 73A. 2003\_CON\_K  
gag.PEP. Fig. 73B. 2003\_CON\_K gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 74A and 74B. Fig. 74A. 2003\_CON\_01\_AE  
5 gag.PEP. Fig. 74B. 2003\_CON\_01\_AE gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 75A and 75B. Fig. 75A. 2003\_CON\_02\_AG  
gag.PEP. Fig. 75B. 2003\_CON\_02\_AG gag.OPT.  
(OPT = codon optimized encoding sequence.)

10 Figures 76A and 76B. Fig. 76A.  
2003\_CON\_03\_ABG gag.PEP. Fig. 76B. 2003\_CON\_03\_ABG  
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 77A and 77B. Fig. 77A.  
2003\_CON\_04\_CFX gag.PEP. Fig. 77B. 2003\_CON\_04\_CFX  
15 gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 78A and 78B. Fig. 78A.  
2003\_CON\_06\_CPX gag.PEP. Fig. 78B. 2003\_CON\_06\_CPX  
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 79A and 79B. Fig. 79A. 2003\_CON\_07\_BC  
20 gag.PEP. Fig. 79B. 2003\_CON\_07\_BC gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 80A and 80B. Fig. 80A. 2003\_CON\_08\_BC  
gag.PEP. Fig. 80B. 2003\_CON\_08\_BC gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 81A and 81B. Fig. 81A. 2003\_CON\_10\_CD  
5 gag.PEP. Fig. 81B. 2003\_CON\_10\_CD gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 82A and 82B. Fig. 82A.  
2003\_CON\_11\_CPX gag.PEP. Fig. 82B. 2003\_CON\_11\_CPX  
gag.OPT. (OPT = codon optimized encoding sequence.)

10 Figures 83A and 83B. Fig. 83A.  
2003\_CON\_12\_BF.gag.PEP. Fig. 83B.  
2003\_CON\_12\_BF.gag.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 84A and 84B. Fig. 84A. 2003\_CON\_14\_BG  
15 gag.PEP. Fig. 84B. 2003\_CON\_14\_BG gag.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 85A and 85B. Fig. 85A. 2003\_CONS  
nef.PEP. Fig. 85B. 2003\_CONS nef.OPT.  
(OPT = codon optimized encoding sequence.)

20 Figures 86A and 86B. Fig. 86A. 2003\_M  
GROUP.anc nef.PEP. Fig. 86B. 2003\_M  
GROUP.anc.nef.OPT. (OPT = codon optimized encoding  
sequence.)

Figures 87A and 87B. Fig. 87A. 2003\_CON\_A  
nef.PEP. Fig. 87B. 2003\_CON\_A nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 88A-88D. Fig. 88A. 2003\_CON\_A1  
5 nef.PEP. Fig. 88B. 2003\_CON\_A1 nef.OPT. Fig. 88C.  
2003\_A1.anc nef.PEP. Fig. 88D. 2003\_A1.anc  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 89A and 89B. Fig. 89A. 2003\_CON\_A2  
nef.PEP. Fig. 89B. 2003\_CON\_A2 nef.OPT.  
10 (OPT = codon optimized encoding sequence.)

Figures 90A-90D. Fig. 90A. 2003\_CON\_B  
nef.PEP. Fig. 90B. 2003\_CON-B nef.OPT. Fig. 90C.  
2003\_B.anc nef.PEP. Fig. 90D. 2003\_B.anc nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 91A and 91B. Fig. 91A. 2003\_CON\_02\_AG  
15 nef.PEP. Fig. 91B. 2003\_CON\_02\_AG nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 92A-92D. Fig. 92A. 2003\_CON\_C  
nef.PEP. Fig. 92B. 2003\_CON\_C nef.OPT. Fig. 92C.  
20 2003\_C.anc nef.PEP. Fig. 92D. 2003\_C.anc nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 93A and 93B. Fig. 93A. 2003\_CON\_D  
nef.PEP. Fig. 93B. 2003\_CON\_D nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 94A and 94B. Fig. 94A. 2003\_CON\_F1  
5 nef.PEP. Fig. 94B. 2003\_CON\_F1 nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 95A and 95B. Fig. 95A. 2003\_CON\_F2  
nef.PEP. Fig. 95B. 2003\_CON\_F2 nef.OPT.  
(OPT = codon optimized encoding sequence.)

10 Figures 96A and 96B. Fig. 96A. 2003\_CON\_G  
nef.PEP. Fig. 96B. 2003\_CON\_G nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 97A and 97B. Fig. 97A. 2003\_CON\_H  
nef.PEP. Fig. 97B. 2003\_CON\_H nef.OPT.  
15 (OPT = codon optimized encoding sequence.)

Figures 98A and 98B. Fig. 98A. 2003\_CON\_01\_AE  
nef.PEP. Fig. 98B. 2003\_CON\_01\_AE nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 99A and 99B. Fig. 99A. 2003\_CON\_03\_AE  
20 nef.PEP. Fig. 99B. 2003\_CON\_03\_AE nef.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 100A and 100B. Fig. 100A.

2003\_CON\_04\_CFX nef.PEP. Fig. 100B.

2003\_CON\_04\_CFX nef.OPT. (OPT = codon optimized  
encoding sequence.)

5 Figures 101A and 101B. Fig. 101A.

2003\_CON\_06\_CFX nef.PEP. Fig. 101B.

2003\_CON\_06\_CFX nef.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 102A and 102B. Fig. 102A.

10 2003\_CON\_08\_BC nef.PEP. Fig. 102B. 2003\_CON\_08\_BC  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 103A and 103B. Fig. 103A.

2003\_CON\_10\_CD nef.PEP. Fig. 103B. 2003\_CON\_10\_CD  
nef.OPT. (OPT = codon optimized encoding sequence.)

15 Figures 104A and 104B. Fig. 104A.

2003\_CON\_11\_CFX nef.PEP. Fig. 104B.

2003\_CON\_11\_CFX nef.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 105A and 105B. Fig. 105A.

20 2003\_CON\_12\_BF nef.PEP. Fig. 105B. 2003\_CON\_12\_BF  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 106A and 106B. Fig. 106A.  
2003\_CON\_14\_BG nef.PEP. Fig. 106B. 2003\_CON\_14\_BG  
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 107A and 107B. Fig. 107A. 2003\_CON\_S  
5 pol.PEP. Fig. 107B. 2003\_CON\_S pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 108A and 108B. Fig. 108A. 2003\_M  
GROUP anc pol.PEP. Fig. 108B. 2003\_M.GROUP anc  
pol.OPT. (OPT = codon optimized encoding sequence.)

10 Figures 109A-109D. Fig. 109A. 2003\_CON\_A1  
pol.PEP. Fig. 109B. 2003\_CON\_A1 pol.OPT.  
Fig. 109C. 2003\_A1.anc pol.PEP. Fig. 109D.  
2003\_A1.anc pol.OPT. (OPT = codon optimized  
encoding sequence.)

15 Figures 110A and 110B. Fig. 110A. 2003\_CON\_A2  
pol.PEP. Fig. 110B. 2003\_CON\_A2 pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 111A-111D. Fig. 111A. 2003\_CON\_B  
pol.PEP. Fig. 111B. 2003\_CON\_B pol.OPT. Fig.  
20 111C. 2003\_B.anc pol.PEP. Fig. 111D. 2003\_B.anc  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 112A-112D. Fig. 112A. 2003\_CON\_C  
pol.PEP. Fig. 112B. 2003\_CON\_C pol.OPT.



Fig. 112C. 2003\_C.anc pol.PEP. Fig. 112D.  
2003\_C.anc pol.OPT. (OPT = codon optimized encoding  
sequence.)

Figures 113A and 113B. Fig. 113A. 2003\_CON\_D  
5 pol.PEP. Fig. 113B. 2003\_CON\_D pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 114A and 114B. Fig. 114A. 2003\_CON\_F1  
pol.PEP. Fig. 114B. 2003\_CON\_F1 pol.OPT.  
(OPT = codon optimized encoding sequence.)

10 Figures 115A and 115B. Fig. 115A. 2003\_CON\_F2  
pol.PEP. Fig. 115B. 2003\_CON\_F2 pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 116A and 116B. Fig. 116A. 2003\_CON\_G  
pol.PEP. Fig. 116B. 2003\_CON\_G pol.OPT.  
15 (OPT = codon optimized encoding sequence.)

Figures 117A and 117B. Fig. 117A. 2003\_CON\_H  
pol.PEP. Fig. 117B. 2003\_CON\_H pol.OPT.  
(OPT = codon optimized encoding sequence.)

Figures 118A and 118B. Fig. 118A.  
20 2003\_CON\_01\_AE pol.PEP. Fig. 118B. 2003\_CON\_01\_AE  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 119A and 119B. Fig. 119A.  
2003\_CON\_02\_AG pol.PEP. Fig. 119B. 2003\_CON\_02\_AG  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 120A and 120B. Fig. 120A.  
5 2003\_CON\_03\_AB pol.PEP. Fig. 120B. 2003\_CON\_03\_AB  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 121A and 121B. Fig. 121A.  
2003\_CON\_04\_CPX pol.PEP. Fig. 121B.  
2003\_CON\_04\_CPX pol.OPT. (OPT = codon optimized  
10 encoding sequence.)

Figures 122A and 122B. Fig. 122A.  
2003\_CON\_06\_CPX pol.PEP. Fig. 122B.  
2003\_CON\_06\_CPX pol.OPT. (OPT = codon optimized  
encoding sequence.)

15 Figures 123A and 123B. Fig. 123A.  
2003\_CON\_08\_BC pol.PEP. Fig. 123B. 2003\_CON\_08\_BC  
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 124A and 124B. Fig. 124A.  
2003\_CON\_10\_CD pol.PEP. Fig. 124B. 2003\_CON\_10\_CD  
20 pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 125A and 125B. Fig. 125A.  
2003\_CON\_11\_CPX pol.PEP. Fig. 125B.

2003\_CON\_11\_CPX pol.OPT. (OPT = codon optimized  
encoding sequence.)

Figures 126A and 126B. Fig. 126A.

2003\_CON\_12\_BF pol.PEP. Fig. 126B. 2003\_CON\_12\_BF  
5 pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 127A and 127B. Fig. 127A.

2003\_CON\_14\_BG pol.PEP. Fig. 127B. 2003\_CON\_14\_BG  
pol.OPT. (OPT = codon optimized encoding sequence.)

#### DETAILED DESCRIPTION OF THE INVENTION

10 The present invention relates to an immunogen  
that induces antibodies that neutralize a wide  
spectrum of human immunodeficiency virus (HIV)  
primary isolates and/or that induces a T cell  
response. The immunogen comprises at least one  
15 consensus or ancestral immunogen (e.g., Env, Gag,  
Nef or Pol), or portion or variant thereof. The  
invention also relates to nucleic acid sequences  
encoding the consensus or ancestral immunogen, or  
portion or variant thereof. The invention further  
20 relates to methods of using both the immunogen and  
the encoding sequences. While the invention is  
described in detail with reference to specific  
consensus and ancestral immunogens (for example, to  
a group M consensus Env), it will be appreciated  
25 that the approach described herein can be used to  
generate a variety of consensus or ancestral

immunogens (for example, envelopes for other HIV-1 groups (e.g., N and O)).

In accordance with one embodiment of the invention, a consensus env gene can be constructed  
5 by generating consensus sequences of env genes for each subtype of a particular HIV-1 group (group M being classified into subtypes A-D, F-H, J and K), for example, from sequences in the Los Alamos HIV Sequence Database (using, for example, MASE  
10 (Multiple Aligned Sequence Editor)). A consensus sequence of all subtype consensus sequences can then be generated to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). In the case  
15 of the group M consensus env gene described in Example 1 (designated CON6), five highly variable regions from a CRF08\_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a region in cytoplasmic domain of gp41) are used to fill in the missing regions in  
20 the sequence (see, however, corresponding regions for Con-S). For high levels of expression, the codons of consensus or ancestral genes can be optimized based on codon usage for highly expressed human genes (Haas et al, Curr. Biol. 6:315-324  
25 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)).

With the Year 1999 consensus group M env gene, CON6, it has been possible to demonstrate induction of superior T cell responses by CON6 versus wild-type B and C env by the number of ELISPOT  
30  $\gamma$ -interferon spleen spot forming cells and the

number of epitopes recognized in two strains of mice (Tables 1 and 2 show the data in BALB/c mice). The ability of CON6 Env protein to induce neutralizing antibodies to HIV-1 primary isolates has been compared to that of several subtype B Env. The target of neutralizing antibodies induced by CON6 includes several non-B HIV-1 strains.

Table 1. T cell epitope mapping of CON6, JRFL and 96ZM651 Env immunogen in BALB/c mice

Peptide	Immunogen			T cell response
	CON6	JRFL (B)	96ZM651 (C)	
CON 6 (group M consensus)				
16 DTEVHNWATHACVP	+		+	CD4
48 KNSSEYYRLINCNTS	+		+	CD4
49 EYYRLINCNTSAITQ				
53 CPKVSFEPIPIHYCA	+			CD4
54 SFEPIPIHYCAPAGF				
62 NVSTVQCTHGKIPVV	+			CD4
104 ETITLPCRIOIINM				
105 LPCRIKQIINMWQGV	+			CD8
130 GIVQQSNLLRAIEA	+			CD4
131 VQOSNLLRAIEAQOHL				
134 AQOHLQLTVWGKIQLO	+			CD4
135 LQLTVWGKIQLOARVL				
Subtype B (MN)				
6223 AKAYDTEVHNWATO	+			CD4
6224 DTEVHNWATOACVP				
6261 ACPKISFEPIPIHYC	+			CD4
6282 ISFEPIPIHYCAPAG				
6286 RKRIHIGPGRAFYTT		+		CD8
6287 HIGPGRAFYTTKNI				
6346 IVQQSNLLRAIEAQ	+			CD4
6347 QNNLLRAIEAQOQHML				
Subtype C (Chn19)				
4834 VPVWKEAKTTLFCASDAKSY			+	CD4
4836 GKEVHNWATHACVPTDNP	+		+	CD4
4848 SSENSSEYYRLINCNTSAIT	+		+	CD4
4854 STVQCTHGKIPVVSTQLLN	+			CD4
4884 QOSNLLRAIEAQOHLQLTV	+			CD4
4885 AQOHLQLTVWGKIQLOTRV	+			CD4

Table 2. T cell epitope mapping of CON6.gp120 immunogen in C57BL/6 mice

Peptide	Peptide sequence	T cell response
CON 6 (consensus)		
2	GIQRNCQHLWRWGT	CD8
3	NCQHLWRWGTMLGM	
16	DTEVHNVWATHACVP	CD4
53	CPKVSFEPIPIHYCA	CD4
97	FYCNTSGLFNSTWMF	CD8
99	FNSTWMFNGTYMFNG	CD8
Subtype B (MN)		
6210	GIRRNQYQHWGWGT	CD8
6211	NYQHWGWGTMLLGL	
6232	NMWKNNMVEQMEDI	CD4
6262	ISFEPIPIHYCAPAG	CD4
6290	NIIGTIRQAHCNISR	CD4
6291	TIRQAHCNISRKWN	
Subtype C (Chn 19)		
4830	MRVTGIRKNYQHLWRWGTM	CD8
5446	RWGTMMLGMLMICSAAEN	CD8
4836	GKEVHNVWATHACVPTDPNP	CD4
4862	GDIRQAHCNISKDKWNETLQ	CD4
4888	LLGIWGCSGKLICTTTVPWN	CD8

For the Year 2000 consensus group M env gene, Con-S, the Con-S envelope has been shown to be as immunogenic as the CON6 envelope gene in T cell  $\gamma$  interferon ELISPOT assays in two strains of mice

(the data for C57BL/6 are shown in Fig. 27).  
Furthermore, in comparing CON6 and Con-S gp140 Envs  
as protein immunogens for antibody in guinea pigs  
(Table 3), both gp140 Envs were found to induce  
5 antibodies that neutralized subtype B primary  
isolates. However, Con-S gp140 also induced robust  
neutralization of the subtype C isolates TV-1 and DU  
123 as well as one subtype A HIV-1 primary isolate,  
while CON6 did not.

TABLE 3 Ability of Group M Consensus CON6 and Con-S Envs to Induce Neutralization of HIV-1 Primary Isolates

HIV-1 Isolate (Subtype)	CON6 gp140CF					CON6 gp140 CFI					CONS gp140 CFI				
	770	771	772	775	781	783	784	786	776	777	778	779	780	781	782
BX08(B)	520	257	428	189	218	164	>540	199	>540	>540	>540	>540	>540	>540	>540
QH0692 (B)	46	55	58	77	<20	91	100	76	109	<20	<20	<20	<20	<20	<20
SS1196(B)	398	306	284	222	431	242	>540	351	>540	296	>540	>540	>540	>540	>540
JRLFL(B)	<20	<20	<20	<20	<20	169	<20	<20	<20	<20	<20	<20	<20	<20	<20
BG1168(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
3988(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
6101(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
TV-1(C)	<20	<20	<20	<20	<20	<20	<20	<20	356	439	>540	>540	>540	>540	>540
DU123(C)	<20	<20	71	74	<20	72	<20	<20	176	329	387	378	378	378	378
DU172(C)	<20	<20	96	64	<20	<20	<20	<20	<20	235	<20	<20	<20	<20	<20
ZM18108.6(C)	ND	ND	ND	ND	<20	<20	<20	<20	84	61	86	43	43	43	43



**≠ 50% Neutralization titers after 4th or 5th immunizations**

Year 2000 Con-S 140CFI<sub>ENV</sub> sequence is shown in Fig. 26A. Gp140 CFI refers to an HIV-1 envelope design in which the cleavage-site is deleted (c), the fusion-site is deleted (F) and the gp41 immunodominant region is deleted (I), in addition to the deletion of transmembrane and cytoplasmic domains. The codon-optimized Year 2000 Con-S 140 CFI sequence is shown in Fig. 26B.

As the next iteration of consensus immunogens, and in recognition of the fact that a practical HIV-1 immunogen can be a polyvalent mixture of either  
5 several subtype consensus genes, a mixture of subtype and consensus genes, or a mixture of centralized genes and wild type genes, a series of 11 subtype consensus, and wild type genes have been designed from subtypes A, B, C, CRF AE01, and G as  
10 well as a group M consensus gene from Year 2003 Los Alamos National Database sequences. The wild type sequences were chosen either because they were known to come from early transmitted HIV-1 strains (those strains most likely to be necessary to be protected  
15 against by a vaccine) or because they were the most recently submitted strains in the database of that subtype. These nucleotide and amino acid sequences are shown in Figures 28-38 (for all 140CF designs shown, 140CF gene can be flanked with the 5'  
20 sequence "TTCAGTCGACGGCCACC" that contains a Kozak sequence (GCCACCATGG/A) and *SalI* site and 3' sequence of TAAAGATCTTACAA containing stop codon and *BglIII* site). Shown in Figures 39-62 are 2003 centralized (consensus and ancestral) HIV-1 envelope  
25 proteins and the codon optimized gene sequences.

Major differences between CON6 gp140 (which does not neutralize non-clade B HIV strains) and Con-S gp140 (which does induce antibodies that neutralize non-clade B HIV strains) are in Con-S V1,  
30 V2, V4 and V5 regions. For clade B strains, peptides of the V3 region can induce neutralizing

antibodies (Haynes et al, J. Immunol. 151:1646-1653 (1993)). Thus, construction of Th-V1, Th-V2, Th-V4, Th-V5 peptides can be expected to give rise to the desired broadly reactive anti-non-clade B neutralizing antibodies. Therefore, the Th-V peptides set forth in Table 4 are contemplated for use as a peptide immunogen(s) derived from Con-S gp140. The gag Th determinant (GTH, Table 4) or any homologous GTH sequence in other HIV strains, can be used to promote immunogenicity and the C4 region of HIV gp120 can be used as well (KQIINMWQVVGKAMYA) or any homologous C4 sequence from other HIV strains (Haynes et al, J. Immunol. 151:1646-1653 (1993)). Con-S V1, V2, V4, V5 peptides with an N-terminal helper determinant can be used singly or together, when formulated in a suitable adjuvant such as Corixa's RC529 (Balldridge et al, J. Endotoxin Res. 8:453-458 (2002)), to induce broadly cross reactive neutralizing antibodies to non-clade B isolates.

20

Table 4		
1)	GTH Con-S V1 132-150	YKRWIILGLNKIVRMYTNVNVTTNTNNTTEEKGEIKN
2)	GTH Con-S V2 157-189	YKRWIILGLNKIVRMYTEIRDKKQKVYALFYRLDVVPIDNNNNSSNYR
3)	GTH Con-S V3 294-315	YKRWIILGLNKIVRMYTRPNNNTRKSIRIGPGQAFYAT
4)	GTH Con-S V4 381-408	YKRWIILGLNKIVRMYNTSGLFNSTWIGNGTKNNNNTNDTITLP
5)	GTH Con-S V5 447-466	YKRWIILGLNKIVRMYRDGGNNNTNETEIFRPGGGD
6)	GTH Con-6 V1 132-150	YKRWIILGLNKIVRMYNVRNVSSNGTETDNEEIKN
7)	GTH Con-6 V2 157-196	YKRWIILGLNKIVRMYTELDRKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYR
8)	GTH-Con6 V3 301-322	YKRWIILGLNKIVRMYTRPNNNTRKSIHIGPGQAFYAT
9)	GTH Con-6 V4 388-418	YKRWIILGLNKIVRMYNTSGLFNSTWMFNGTYMFNGTKDNSETITLP
10)	GTH Con 6 V5 457-477	YKRWIILGLNKIVRMYRDGGNNSNKNKTETFRPGGGD

It will be appreciated that the invention includes portions and variants of the sequences specifically disclosed herein. For example, forms of codon optimized consensus encoding sequences can be constructed as gp140CF, gp140 CFI, gp120 or gp160 forms with either gp120/41 cleaved or uncleaved. For example, and as regards the consensus and ancestral envelope sequences, the invention encompasses envelope sequences devoid of V3. Alternatively, V3 sequences can be selected from preferred sequences, for example, those described in U.S. Application No. 10/431,596 and U.S. Provisional Application No. 60/471,327. In addition, an optimal immunogen for breadth of response can include mixtures of group M consensus *gag*, *pol*, *nef* and *env* encoding sequences, and as well as consist of

mixtures of subtype consensus or ancestral encoding sequences for *gag*, *pol*, *nef* and *env* HIV genes. For dealing with regional differences in virus strains, an efficacious mixture can include mixtures of  
5 consensus/ancestral and wild type encoding sequences.

A consensus or ancestral envelope of the invention can be been "activated" to expose intermediate conformations of neutralization  
10 epitopes that normally are only transiently or less well exposed on the surface of the HIV virion. The immunogen can be a "frozen" triggered form of a consensus or ancestral envelope that makes available specific epitopes for presentation to B lymphocytes.  
15 The result of this epitope presentation is the production of antibodies that broadly neutralize HIV. (Attention is directed to WO 02/024149 and to the activated/triggered envelopes described therein.)

20 The concept of a fusion intermediate immunogen is consistent with observations that the gp41 HR-2 region peptide, DP178, can capture an uncoiled conformation of gp41 (Furata et al, Nature Struct. Biol. 5:276 (1998)), and that formalin-fixed HIV-  
25 infected cells can generate broadly neutralizing antibodies (LaCasse et al, Science 283:357 (1997)). Recently a monoclonal antibody against the coiled-coil region bound to a conformational determinant of gp41 in HR1 and HR2 regions of the coiled-coil gp41  
30 structure, but did not neutralize HIV (Jiang et al, J. Virol. 10213 (1998)). However, this latter study

proved that the coiled-coil region is available for antibody to bind if the correct antibody is generated.

The immunogen of one aspect of the invention  
5 comprises a consensus or ancestral envelope either in soluble form or anchored, for example, in cell vesicles or in liposomes containing translipid bilayer envelope. To make a more native envelope, gp140 or gp160 consensus or ancestral sequences can  
10 be configured in lipid bilayers for native trimeric envelope formation. Alternatively, triggered gp160 in aldrithio 1-2 inactivated HIV-1 virions can be used as an immunogen. The gp160 can also exist as a recombinant protein either as gp160 or gp140 (gp140  
15 is gp160 with the transmembrane region and possibly other gp41 regions deleted). Bound to gp160 or gp140 can be recombinant CCR5 or CXCR4 co-receptor proteins (or their extracellular domain peptide or protein fragments) or antibodies or other ligands  
20 that bind to the CXCR4 or CCR5 binding site on gp120, and/or soluble CD4, or antibodies or other ligands that mimic the binding actions of CD4. Alternatively, vesicles or liposomes containing CD4, CCR5 (or CXCR4), or soluble CD4 and peptides  
25 reflective of CCR5 or CXCR4 gp120 binding sites. Alternatively, an optimal CCR5 peptide ligand can be a peptide from the N-terminus of CCR5 wherein specific tyrosines are sulfated (Bormier et al, Proc. Natl. Acad. Sci. USA 97:5762 (2001)). The  
30 triggered immunogen may not need to be bound to a membrane but may exist and be triggered in solution.

Alternatively, soluble CD4 (sCD4) can be replaced by an envelope (gp140 or gp160) triggered by CD4 peptide mimetopes (Vitra et al, Proc. Natl. Acad. Sci. USA 96:1301 (1999)). Other HIV co-receptor molecules that "trigger" the gp160 or gp140 to undergo changes associated with a structure of gp160 that induces cell fusion can also be used. Ligation of soluble HIV gp140 primary isolate HIV 89.6 envelope with soluble CD4 (sCD4) induced conformational changes in gp41.

In one embodiment, the invention relates to an immunogen that has the characteristics of a receptor (CD4)-ligated consensus or ancestral envelope with CCR5 binding region exposed but unlike CD4-ligated proteins that have the CD4 binding site blocked, this immunogen has the CD4 binding site exposed (open). Moreover, this immunogen can be devoid of host CD4, which avoids the production of potentially harmful anti-CD4 antibodies upon administration to a host.

The immunogen can comprise consensus or ancestral envelope ligated with a ligand that binds to a site on gp120 recognized by an A32 monoclonal antibodies (mab) (Wyatt et al, J. Virol. 69:5723 (1995), Boots et al, AIDS Res. Hum. Retro. 13:1549 (1997), Moore et al, J. Virol. 68:8350 (1994), Sullivan et al, J. Virol. 72:4694 (1998), Fouts et al, J. Virol. 71:2779 (1997), Ye et al, J. Virol. 74:11955 (2000)). One A32 mab has been shown to mimic CD4 and when bound to gp120, upregulates (exposes) the CCR5 binding site (Wyatt et al, J.

Virology 69:5723 (1995)). Ligation of gp120 with such a ligand also upregulates the CD4 binding site and does not block CD4 binding to gp120.

Advantageously, such ligands also upregulate the HR-2 binding site of gp41 bound to cleaved gp120, uncleaved gp140 and cleaved gp41, thereby further exposing HR-2 binding sites on these proteins - each of which are potential targets for anti-HIV neutralizing antibodies.

10 In a specific aspect of this embodiment, the immunogen comprises soluble HIV consensus or ancestral gp120 envelope ligated with either an intact A32 mab, a Fab2 fragment of an A32 mab, or a Fab fragment of an A32 mab, with the result that the  
15 CD4 binding site, the CCR5 binding site and the HR-2 binding site on the consensus or ancestral envelope are exposed/upregulated. The immunogen can comprise consensus or ancestral envelope with an A32 mab (or fragment thereof) bound or can comprise consensus or  
20 ancestral envelope with an A32 mab (or fragment thereof) bound and cross-linked with a cross-linker such as .3% formaldehyde or a heterobifunctional cross-linker such as DTSSP (Pierce Chemical Company). The immunogen can also comprise uncleaved  
25 consensus or ancestral gp140 or a mixture of uncleaved gp140, cleaved gp41 and cleaved gp120. An A32 mab (or fragment thereof) bound to consensus or ancestral gp140 and/or gp120 or to gp120 non-covalently bound to gp41, results in upregulation  
30 (exposure) of HR-2 binding sites in gp41, gp120 and uncleaved gp140. Binding of an A32 mab (or fragment



thereof) to gp120 or gp140 also results in upregulation of the CD4 binding site and the CCR5 binding site. As with gp120 containing complexes, complexes comprising uncleaved gp140 and an A32 mab  
5 (or fragment thereof) can be used as an immunogen uncross-linked or cross-linked with cross-linker such as .3% formaldehyde or DTSSP. In one embodiment, the invention relates to an immunogen comprising soluble uncleaved consensus or ancestral  
10 gp140 bound and cross linked to a Fab fragment or whole A32 mab, optionally bound and cross-linked to an HR-2 binding protein.

The consensus or ancestral envelope protein triggered with a ligand that binds to the A32 mab  
15 binding site on gp120 can be administered in combination with at least a second immunogen comprising a second envelope, triggered by a ligand that binds to a site distinct from the A32 mab binding site, such as the CCR5 binding site  
20 recognized by mab 17b. The 17b mab (Kwong et al, Nature 393:648 (1998) available from the AIDS Reference Repository, NIAID, NIH) augments sCD4 binding to gp120. This second immunogen (which can also be used alone or in combination with triggered  
25 immunogens other than that described above) can, for example, comprise soluble HIV consensus or ancestral envelope ligated with either the whole 17b mab, a Fab2 fragment of the 17b mab, or a Fab fragment of the 17b mab. It will be appreciated that other CCR5  
30 ligands, including other antibodies (or fragments thereof), that result in the CD4 binding site being

exposed can be used in lieu of the 17b mab. This further immunogen can comprise gp120 with the 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) bound or can comprise gp120 with  
5 the 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) bound and cross-linked with an agent such as .3% formaldehyde or a heterobifunctional cross-linker, such as DTSSP (Pierce Chemical Company). Alternatively, this  
10 further immunogen can comprise uncleaved gp140 present alone or in a mixture of cleaved gp41 and cleaved gp120. Mab 17b, or fragment thereof (or other CCR5 ligand as indicated above) bound to gp140 and/or gp120 in such a mixture results in exposure  
15 of the CD4 binding region. The 17b mab, or fragment thereof, (or other CCR5 ligand as indicated above) gp140 complexes can be present uncross-linked or cross-linked with an agent such as .3% formaldehyde or DTSSP.

20 Soluble HR-2 peptides, such as T649Q26L and DP178, can be added to the above-described complexes to stabilize epitopes on consensus gp120 and gp41 as well as uncleaved consensus gp140 molecules, and can be administered either cross-linked or uncross-  
25 linked with the complex.

A series of monoclonal antibodies (mabs) have been made that neutralize many HIV primary isolates, including, in addition to the 17b mab described above, mab IgG1b12 that binds to the CD4 binding  
30 site on gp120 (Roben et al, J. Virol. 68:482 (1994), Mo et al, J. Virol. 71:6869 (1997)), mab 2G12 that

binds to a conformational determinant on gp120 (Trkola et al, J. Virol. 70:1100 (1996)), and mab 2F5 that binds to a membrane proximal region of gp41 (Muster et al, J. Virol. 68:4031 (1994)).

5       As indicated above, various approaches can be used to "freeze" fusogenic epitopes in accordance with the invention. For example, "freezing" can be effected by addition of the DP-178 or T-649Q26L peptides that represent portions of the coiled coil  
10 region, and that when added to CD4-triggered consensus or ancestral envelope, result in prevention of fusion (Rimsky et al, J. Virol. 72:986-993 (1998)). HR-2 peptide bound consensus or ancestral gp120, gp140, gp41 or gp160 can be used as  
15 an immunogen or crosslinked by a reagent such as DTSSP or DSP (Pierce Co.), formaldehyde or other crosslinking agent that has a similar effect.

"Freezing" can also be effected by the addition of 0.1% to 3% formaldehyde or paraformaldehyde, both  
20 protein cross-linking agents, to the complex, to stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160 complex, or to stabilize the "triggered" gp41 molecule, or both (LaCasse et al, Science 283:357-362 (1999)).

25       Further, "freezing" of consensus or ancestral gp41 or gp120 fusion intermediates can be effected by addition of heterobifunctional agents such as DSP (dithiobis[succinimidylpropionate]) (Pierce Co. Rockford, ILL., No. 22585ZZ) or the water soluble  
30 DTSSP (Pierce Co.) that use two NHS esters that are reactive with amino groups to cross link and

stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160 complex, or to stabilize the "triggered" gp41 molecule, or both.

Analysis of T cell immune responses in  
5 immunized or vaccinated animals and humans shows that the envelope protein is normally not a main target for T cell immune response although it is the only gene that induces neutralizing antibodies. HIV-1 Gag, Pol and Nef proteins induce a potent T  
10 cell immune response. Accordingly, the invention includes a repertoire of consensus or ancestral immunogens that can induce both humoral and cellular immune responses. Subunits of consensus or ancestral sequences can be used as T or B cell  
15 immunogens. (See Examples 6 and 7, and Figures referenced therein, and Figures 63-127.

The immunogen of the invention can be formulated with a pharmaceutically acceptable carrier and/or adjuvant (such as alum) using  
20 techniques well known in the art. Suitable routes of administration of the present immunogen include systemic (e.g. intramuscular or subcutaneous). Alternative routes can be used when an immune response is sought in a mucosal immune system (e.g.,  
25 intranasal).

The immunogens of the invention can be chemically synthesized and purified using methods which are well known to the ordinarily skilled artisan. The immunogens can also be synthesized by  
30 well-known recombinant DNA techniques. Nucleic acids encoding the immunogens of the invention can

be used as components of, for example, a DNA vaccine wherein the encoding sequence is administered as naked DNA or, for example, a minigene encoding the immunogen can be present in a viral vector. The  
5 encoding sequence can be present, for example, in a replicating or non-replicating adenoviral vector, an adeno-associated virus vector, an attenuated mycobacterium tuberculosis vector, a Bacillus Calmette Guerin (BCG) vector, a vaccinia or Modified  
10 Vaccinia Ankara (MVA) vector, another pox virus vector, recombinant polio and other enteric virus vector, Salmonella species bacterial vector, Shigella species bacterial vector, Venezuelan Equine Encephalitis Virus (VEE) vector, a Semliki  
15 Forest Virus vector, or a Tobacco Mosaic Virus vector. The encoding sequence, can also be expressed as a DNA plasmid with, for example, an active promoter such as a CMV promoter. Other live vectors can also be used to express the sequences of  
20 the invention. Expression of the immunogen of the invention can be induced in a patient's own cells, by introduction into those cells of nucleic acids that encode the immunogen, preferably using codons and promoters that optimize expression in human  
25 cells. Examples of methods of making and using DNA vaccines are disclosed in U.S. Pat. Nos. 5,580,859, 5,589,466, and 5,703,055.

The composition of the invention comprises an immunologically effective amount of the immunogen of  
30 this invention, or nucleic acid sequence encoding same, in a pharmaceutically acceptable delivery

system. The compositions can be used for prevention and/or treatment of immunodeficiency virus infection. The compositions of the invention can be formulated using adjuvants, emulsifiers,  
5 pharmaceutically-acceptable carriers or other ingredients routinely provided in vaccine compositions. Optimum formulations can be readily designed by one of ordinary skill in the art and can include formulations for immediate release and/or  
10 for sustained release, and for induction of systemic immunity and/or induction of localized mucosal immunity (e.g, the formulation can be designed for intranasal administration). The present compositions can be administered by any convenient  
15 route including subcutaneous, intranasal, oral, intramuscular, or other parenteral or enteral route. The immunogens can be administered as a single dose or multiple doses. Optimum immunization schedules can be readily determined by the ordinarily skilled  
20 artisan and can vary with the patient, the composition and the effect sought.

The invention contemplates the direct use of both the immunogen of the invention and/or nucleic acids encoding same and/or the immunogen expressed  
25 as minigenes in the vectors indicated above. For example, a minigene encoding the immunogen can be used as a prime and/or boost.

The invention includes any and all amino acid sequences disclosed herein and, where applicable, CF  
30 and CFI forms thereof, as well as nucleic acid

sequences encoding same (and nucleic acids complementary to such encoding sequences).

Certain aspects of the invention can be described in greater detail in the non-limiting  
5 Examples that follows.

#### EXAMPLE 1

##### Artificial HIV-1 Group M Consensus Envelope

#### EXPERIMENTAL DETAILS

10       *Expression of CON6 gp120 and gp140 proteins in recombinant vaccinia viruses (VV).* To express and purify the secreted form of HIV-1 CON6 envelope proteins, CON6 gp120 and gp140CF plasmids were constructed by introducing stop codons after the  
15 gp120 cleavage site (REKR) and before the transmembrane domain (YIKIFIMIVGGLIGLRIVFAVLSIVN), respectively. The gp120/gp41 cleavage site and fusion domain of gp41 were deleted in the gp140CF protein. Both CON6 gp120 and gp140CF DNA constructs  
20 were cloned into the pSC65 vector (from Bernard Moss, NIH, Bethesda, MD) at SalI and KpnI restriction enzyme sites. This vector contains the lacZ gene that is controlled by the p7.5 promoter. A back-to-back P E/L promoter was used to express  
25 CON6 env genes. BSC-1 cells were seeded at  $2 \times 10^5$  in each well in a 6-well plate, infected with wild-type vaccinia virus (WR) at a MOI of 0.1 pfu/cell, and 2 hr after infection, pSC65-derived plasmids

containing CON6 *env* genes were transfected into the VV-infected cells and recombinant (r) VV selected as described (Moss and Earl, Current Protocols in Molecular Biology, eds, Ausubel et al (John Wiley & Sons, Inc. Indianapolis, IN) pp. 16.15.1-16.19.9 (1998)). Recombinant VV that contained the CON6 *env* genes were confirmed by PCR and sequencing analysis. Expression of the CON6 envelope proteins was confirmed by SDS-PAGE and Western blot assay.

10 Recombinant CON6 gp120 and gp140CF were purified with agarose *galanthus Nivalis* lectin beads (Vector Labs, Burlingame, CA), and stored at -70°C until use. Recombinant VV expressing JRFL (vCB-28) or 96ZM651 (vT241R) gp160 were obtained from the NIH AIDS

15 Research and Reference Reagent Program (Bethesda, MD).

*Monoclonal Antibodies and gp120 Wild-type Envelopes.* Human mabs against a conformational

20 determinant on gp120 (A32), the gp120 V3 loop (F39F) and the CCR5 binding site (17b) were the gifts of James Robinson (Tulane Medical School, New Orleans, LA) (Wyatt et al, Nature 393:705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). Mabs 2F5,

25 447, b12, 2G12 and soluble CD4 were obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD) (Gorny et al, J. Immunol. 159:5114-5122 (1997), Nyambi et al, J. Virol. 70:6235-6243 (1996), Purtscher et al, AIDS Res. Hum. Retroviruses

30 10:1651-1658 (1994), Trkola et al, J. Virol 70:1100-1108 (1996)). T8 is a murine mab that maps to the



gp120 C1 region (a gift from P. Earl, NIH, Bethesda, MD). BaL (subtype B), 96ZM651 (subtype C), and 93TH975 (subtype E) gp120s were provided by QBI, Inc. and the Division of AIDS, NIH. CHO cell lines  
5 that express 92U037 (subtype A) and 93BR029 (subtype F) gp140 (secreted and uncleaved) were obtained from NICBS, England.

*Surface Plasmon Resonance Biosensor (SPR)*

10 *Measurements and ELISA.* SPR biosensor measurements were determined on a BIAcore 3000 instrument (BIAcore Inc., Uppsala, Sweden) instrument and data analysis was performed using BIAevaluation 3.0 software (BIAcore Inc, Uppsala, Sweden). Anti-gp120  
15 mabs (T8, A32, 17b, 2G12) or sCD4 in 10mM Na-acetate buffer, pH 4.5 were directly immobilized to a CM5 sensor chip using a standard amine coupling protocol for protein immobilization. FPLC purified CON6 gp120 monomer or gp140CF oligomer recombinant  
20 proteins were flowed over CM5 sensor chips at concentrations of 100 and 300 µg/ml, respectively. A blank in-line reference surface (activated and de-activated for amine coupling) or non-bonding mab controls were used to subtract non-specific or bulk  
25 responses. Soluble 89.6 gp120 and irrelevant IgG was used as a positive and negative control respectively and to ensure activity of each mab surface prior to injecting the CON6 Env proteins. Binding of CON6 envelope proteins was monitored in  
30 real-time at 25°C with a continuous flow of PBS (150 mM NaCl, 0.005% surfactant P20), pH 7.4 at 10-30

μl/min. Bound proteins were removed and the sensor surfaces were regenerated following each cycle of binding by single or duplicate 5-10 μl pulses of regeneration solution (10 mM glycine-HCl, pH 2.9).

5 ELISA was performed to determine the reactivity of various mabs to CON6 gp120 and gp140CF proteins as described (Haynes et al, AIDS Res. Hum. Retroviruses 11:211-221 (1995)). For assay of human mab binding to rgp120 or gp140 proteins, end-point titers were  
10 defined as the highest titer of mab (beginning at 20 μg/ml) at which the mab bound CON6 gp120 and gp140CF Env proteins ≥ 3 fold over background control (non-binding human mab).

15 *Infectivity and coreceptor usage assays.* HIV-1/SG3Δenv and CON6 or control env plasmids were cotransfected into human 293T cells. Pseudotyped viruses were harvested, filtered and p24 concentration was quantitated (DuPont/NEN Life  
20 Sciences, Boston, MA). Equal amounts of p24 (5 ng) for each pseudovirion were used to infect JC53-BL cells to determine the infectivity (Derdeyn et al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). JC53-BL  
25 cells express CD4, CCR5 and CXCR4 receptors and contain a β-galactosidase (β-gal) gene stably integrated under the transcriptional control of an HIV-1 long terminal repeat (LTR). These cells can be used to quantify the infectious titers of  
30 pseudovirion stocks by staining for β-gal expression

and counting the number of blue cells (infectious units) per microgram of p24 of pseudovirions (IU/ $\mu$ g p24) (Derdeyn et al, J. Virol. 74:8358-8367 (2000), Wei et al, Antimicrob Agents Chemother. 46:1896-1905 (2002)). To determine the coreceptor usage of the CON6 env gene, JC53BL cells were treated with 1.2  $\mu$ M AMD3100 and 4  $\mu$ M TAK-799 for 1 hr at 37°C then infected with equal amounts of p24 (5 ng) of each Env pseudotyped virus. The blockage efficiency was expressed as the percentage of the infectious units from blockage experiments compared to that from control culture without blocking agents. The infectivity from control group (no blocking agent) was arbitrarily set as 100%.

15

*Immunizations.* All animals were housed in the Duke University Animal Facility under AALAC guidelines with animal use protocols approved by the Duke University Animal Use and Care Committee. Recombinant CON6 gp120 and gp140CF glycoproteins were formulated in a stable emulsion with RIBI-CWS adjuvant based on the protocol provided by the manufacturer (Sigma Chemical Co., St. Louis, MO). For induction of anti-envelope antibodies, each of four out-bred guinea pigs (Harlan Sprague, Inc., Chicago, IL) was given 100  $\mu$ g either purified CON6 gp120 or gp140CF subcutaneously every 3 weeks (total of 5 immunizations). Serum samples were heat-inactivated (56°C, 1 hr), and stored at -20°C until use.

30

For induction of anti-envelope T cell responses, 6-8 wk old female BALB/c mice (Frederick Cancer Research and Developmental Center, NCI, Frederick, MD) were immunized i.m. in the quadriceps  
5 with 50  $\mu$ g plasmid DNA three times at a 3-week interval. Three weeks after the last DNA immunization, mice were boosted with  $10^7$  PFU of rVV expressing Env proteins. Two weeks after the boost, all mice were euthanized and spleens were removed  
10 for isolation of splenocytes.

*Neutralization assays.* Neutralization assays were performed using either a MT-2 assay as described in Bures et al, AIDS Res. Hum.  
15 Retroviruses 16:2019-2035 (2000), a luciferase-based multiple replication cycle HIV-1 infectivity assay in 5.25.GFP.Luc.M7 cells using a panel of HIV-1 primary isolates (Bures et al, AIDS Res. Hum. Retroviruses 16:2019-2035 (2000), Bures et al, J.  
20 Virol. 76:2233-2244 (2002)), or a syncytium (fusion from without) inhibition assay using inactivated HIV-1 virions (Rossio et al, J. Virol. 72:7992-8001 (1998)). In the luciferase-based assay, neutralizing antibodies were measured as a function  
25 of a reduction in luciferase activity in 5.25.EGFP.Luc.M7 cells provided by Nathaniel R. Landau, Salk Institute, La Jolla, CA (Brandt et al, J. Biol. Chem. 277:17291-17299 (2002)). Five  
hundred tissue culture infectious dose 50 (TCID<sub>50</sub>) of  
30 cell-free virus was incubated with indicated serum

dilutions in 150  $\mu$ l (1 hr, at 37°C) in triplicate in 96-well flat-bottom culture plates. The 5.25.EGFP.Luc.M7 cells were suspended at a density of  $5 \times 10^5$ /ml in media containing DEAE dextran (10  $\mu$ g/ml). Cells (100  $\mu$ l) were added and until 10% of cells in control wells (no test serum sample) were positive for GFP expression by fluorescence microscopy. At this time the cells were concentrated 2-fold by removing one-half volume of media. A 50  $\mu$ l suspension of cells was transferred to 96-well white solid plates (Costar, Cambridge, MA) for measurement of luciferase activity using Bright-Glo™ substrate (Promega, Madison, WI) on a Wallac 1420 Multilabel Counter (PerkinElmer Life Sciences, Boston, MA). Neutralization titers in the MT-2 and luciferase assays were those where  $\geq 50\%$  virus infection was inhibited. Only values that titered beyond 1:20 (i.e.  $>1:30$ ) were considered significantly positive. The syncytium inhibition "fusion from without" assay utilized HIV-1 aldrithiol-2 (AT-2) inactivated virions from HIV-1 subtype B strains ADA and AD8 (the gift of Larry Arthur and Jeffrey Lifson, Frederick Research Cancer Facility, Frederick, MD) added to SupT1 cells, with syncytium inhibition titers determined as those titers where  $\geq 90\%$  of syncytia were inhibited compared to prebleed sera.

*Enzyme linked immune spot (ELISPOT) assay.*

Single-cell suspensions of splenocytes from

individual immunized mice were prepared by mincing and forcing through a 70  $\mu$ m Nylon cell strainer (BD Labware, Franklin Lakes, NJ). Overlapping Env peptides of CON6 gp140 (159 peptides, 15mers overlapping by 11) were purchased from Boston Bioscience, Inc (Royal Oak, MI). Overlapping Env peptides of MN gp140 (subtype B; 170 peptides, 15mers overlapping by 11) and Chn19 gp140 (subtype C; 69 peptides, 20mers overlapping by 10) were obtained from the NIH AIDS Research and Reference Reagent Program (Bethesda, MD). Splenocytes (5 mice/group) from each mouse were stimulated *in vitro* with overlapping Env peptides pools from CON6, subtype B and subtype C Env proteins. 96-well PVDF plates (MultiScreen-IP, Millipore, Billerica, MA) were coated with anti-IFN- $\gamma$  mab (5  $\mu$ g/ml, AN18; Mabtech, Stockholm, Sweden). After the plates were blocked at 37°C for 2 hr using complete Hepes buffered RPMI medium, 50 $\mu$ l of the pooled overlapping envelope peptides (13 CON6 and MN pools, 13-14 peptides in each pool; 9 Chn19 pool, 7-8 peptide in each pool) at a final concentration of 5  $\mu$ g/ml of each were added to the plate. Then 50  $\mu$ l of splenocytes at a concentration of  $1.0 \times 10^7$ /ml were added to the wells in duplicate and incubated for 16 hr at 37°C with 5% CO<sub>2</sub>. The plates were incubated with 100  $\mu$ l of a 1:1000 dilution of streptavidin alkaline phosphatase (Mabtech, Stockholm, Sweden), and purple spots developed using 100  $\mu$ l of BCIP/NBT (Plus) Alkaline Phosphatase Substrate (Moss,

Pasadena, MD). Spot forming cells (SFC) were measured using an Immunospot counting system (CTL Analyzers, Cleveland, OH). Total responses for each envelope peptide pool are expressed as SFCs per 10<sup>6</sup> splenocytes.

## RESULTS

*CON6 Envelope Gene Design, Construction and Expression.* An artificial group M consensus env gene (CON6) was constructed by generating consensus sequences of env genes for each HIV-1 subtype from sequences in the Los Alamos HIV Sequence Database, and then generating a consensus sequence of all subtype consensus sequences to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). Five highly variable regions from a CRF08\_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a region in cytoplasmic domain of gp41) were then used to fill in the missing regions in CON6 sequence. The CON6 V3 region is group M consensus (Figure 1A). For high levels of expression, the codons of CON6 env gene were optimized based on codon usage for highly expressed human genes (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)). (See Fig. 1D.) The codon optimized CON6 env gene was constructed and subcloned into pCDNA3.1 DNA at EcoR I and BamH I sites (Gao et al, AIDS Res. Hum. Retroviruses, 19:817-823 (2003)). High levels of protein

expression were confirmed with Western-blot assays after transfection into 293T cells. To obtain recombinant CON6 Env proteins for characterization and use as immunogens, rVV was generated to express  
5 secreted gp120 and uncleaved gp140CF (Figure 1B). Purity for each protein was  $\geq 90\%$  as determined by Coomassie blue gels under reducing conditions (Figure 1C).

10 *CD4 Binding Domain and Other Wild-type HIV-1 Epitopes are Preserved on CON6 Proteins.* To determine if CON6 proteins can bind to CD4 and express other wild-type HIV-1 epitopes, the ability of CON6 gp120 and gp140CF to bind soluble(s) CD4, to  
15 bind several well-characterized anti-gp120 mabs, and to undergo CD4-induced conformational changes was assayed. First, BIAcore CM5 sensor chips were coated with either sCD4 or mabs to monitor their binding activity to CON6 Env proteins. It was found  
20 that both monomeric CON6 gp120 and oligomeric gp140CF efficiently bound sCD4 and anti-gp120 mabs T8, 2G12 and A32, but did not constitutively bind mab 17b, that recognizes a CD4 inducible epitope in the CCR5 binding site of gp120 (Figures 2A and 2B).  
25 Both sCD4 and A32 can expose the 17b binding epitope after binding to wild-type gp120 (Wyatt et al, Nature 393;705-711 (1998), Wyatt et al, J. Virol. 69:5723-5733 (1995)). To determine if the 17b epitope could be induced on CON6 Envs by either sCD4  
30 or A32, sCD4, A32 and T8 were coated on sensor chips, then CON6 gp120 or gp140CF captured, and mab



17b binding activity monitored. After binding sCD4 or mab A32, both CON6 gp120 and gp140CF were triggered to undergo conformational changes and bound mab 17b (Figures 2C and 2D). In contrast, after binding mab T8, the 17b epitope was not exposed (Figures 2C and 2D). ELISA was next used to determine the reactivity of a panel of human mabs against the gp120 V3 loop (447, F39F), the CD4 binding site (b12), and the gp41 neutralizing determinant (2F5) to CON6 gp120 and gp140CF (Figure 2E). Both CON6 rgp120 and rgp140CF proteins bound well to neutralizing V3 mabs 447 and F39F and to the potent neutralizing CD4 binding site mab b12. Mab 2F5, that neutralizes HIV-1 primary isolates by binding to a C-terminal gp41 epitope, also bound well to CON6 gp140CF (Figure 2E).

*CON6 env Gene is Biologically Functional and Uses CCR5 as its Coreceptor.* To determine whether CON6 envelope gene is biologically functional, it was co-transfected with the env-defective SG3 proviral clone into 293T cells. The pseudotyped viruses were harvested and JC53BL cells infected. Blue cells were detected in JC53-BL cells infected with the CON6 Env pseudovirions, suggesting that CON6 Env protein is biologically functional (Figure 3A). However, the infectious titers were 1-2 logs lower than that of pseudovirions with either YU2 or NL4-3 wild-type HIV-1 envelopes.

The co-receptor usage for the CON6 env gene was next determined. When treated with CXCR4 blocking

agent AMD3100, the infectivity of NL4-3 Env-pseudovirions was blocked while the infectivity of YU2 or CON6 Env-pseudovirions was not inhibited (Figure 3B). In contrast, when treated with CCR5 blocking agent TAK-779, the infectivity of NL4-3 Env-pseudovirions was not affected, while the infectivity of YU2 or CON6 Env-pseudovirions was inhibited. When treated with both blocking agents, the infectivity of all pseudovirions was inhibited. Taken together, these data show that the CON6 envelope uses the CCR5 co-receptor for its entry into target cells.

*Reaction of CON6 gp120 With Different Subtype Sera.* To determine if multiple subtype linear epitopes are preserved on CON6 gp120, a recombinant Env protein panel (gp120 and gp140) was generated. Equal amounts of each Env protein (100 ng) were loaded on SDS-polyacrylamide gels, transferred to nitrocellulose, and reacted with subtype A through G patient sera as well as anti-CON6 gp120 guinea pig sera (1:1,000 dilution) in Western blot assays. For each HIV-1 subtype, four to six patient sera were tested. One serum representative for each subtype is shown in Figure 4.

It was found that whereas all subtype sera tested showed variable reactivities among Envs in the panel, all group M subtype patient sera reacted equally well with CON6 gp120 Env protein, demonstrating that wild-type HIV-1 Env epitopes recognized by patient sera were well preserved on

the CON6 Env protein. A test was next made as to whether CON6 gp120 antiserum raised in guinea pigs could react to different subtype Env proteins. It was found that the CON6 serum reacted to its own and  
5 other subtype Env proteins equally well, with the exception of subtype A Env protein (Figure 4).

*Induction of T Cell Responses to CON6, Subtype B and Subtype C Envelope Overlapping Peptides.* To  
10 compare T cell immune responses induced by CON6 Env immunogens with those induced by subtype specific immunogens, two additional groups of mice were immunized with subtype B or subtype C DNAs and with corresponding rVV expressing subtype B or C envelope  
15 proteins. Mice immunized with subtype B (JRFL) or subtype C (96ZM651) Env immunogen had primarily subtype-specific T cell immune responses (Figure 5). IFN- $\gamma$  SFCs from mice immunized with JRFL (subtype B) immunogen were detected after stimulation with  
20 subtype B (MN) peptide pools, but not with either subtype C (Chn19) or CON6 peptide pools. IFN- $\gamma$  SFCs from mice immunized with 96ZM651 (subtype C) immunogen were detected after the stimulation with  
25 both subtype C (Chn19) and CON6 peptide pools, but not with subtype B (MN) peptide pools. In contrast, IFN- $\gamma$  SFCs were identified from mice immunized with CON6 Env immunogens when stimulated with either CON6 peptide pools as well as by subtype B or C peptide  
30 pools (Figure 5). The T cell immune responses induced by CON6 gp140 appeared more robust than

those induced by CON6 gp120. Taken together, these data demonstrated that CON6 gp120 and gp140CF immunogens were capable of inducing T cell responses that recognized T cell epitopes of wild-type subtype B and C envelopes.

*Induction of Antibodies by Recombinant CON6 gp120 and gp140CF Envelopes that Neutralize HIV-1 Subtype B and C Primary Isolates.* To determine if the CON6 envelope immunogens can induce antibodies that neutralize HIV-1 primary isolates, guinea pigs were immunized with either CON6 gp120 or gp140CF protein. Sera collected after 4 or 5 immunizations were used for neutralization assays and compared to the corresponding prebleed sera. Two AT-2 inactivated HIV-1 isolates (ADA and AD8) were tested in syncytium inhibition assays (Table 5A). Two subtype B SHIV isolates, eight subtype B primary isolates, four subtype C, and one each subtype A, D, and E primary isolates were tested in either the MT-2 or the luciferase-based assay (Table 5B). In the syncytium inhibition assay, it was found that antibodies induced by both CON 6 gp120 and gp140CF proteins strongly inhibited AT-2 inactivated ADA and AD8-induced syncytia (Table 5A). In the MT-2 assay, weak neutralization of 1 of 2 SHIV isolates (SHIV SF162P3) by two gp120 and one gp140CF sera was found (Table 5B). In the luciferase-based assay, strong neutralization of 4 of 8 subtype B primary isolates (BX08, SF162, SS1196, and BAL) by all gp120 and gp140CF sera was found, and weak neutralization of 2

- of 8 subtype B isolates (6101, 0692) by most gp120 and gp140CF sera was found. No neutralization was detected against HIV-1 PAVO (Table 5B). Next, the CON6 anti-gp120 and gp140CF sera were tested against
- 5 four subtype C HIV-1 isolates, and weak neutralization of 3 of 4 isolates (DU179, DU368, and S080) was found, primarily by anti-CON6 gp120 sera. One gp140CF serum, no. 653, strongly neutralized DU179 and weakly neutralized S080 (Table 5B).
- 10 Finally, anti-CON6 Env sera strongly neutralized a subtype D isolate (93ZR001), weakly neutralized a subtype E (CM244) isolate, and did not neutralize a subtype A (92RW020) isolate.

Table 5A

**Ability of HIV-1 Group M Consensus Envelope CON6 Proteins to Induce Fusion Inhibiting Antibodies**

Guinea Pig No.	Immunogen	Syncytium Inhibition antibody titer <sup>1</sup>	
		AD8	ADA
646	gp120	270	270
647	gp120	90	90
648	gp120	90	270
649	gp120	90	90
Geometric Mean Titer		119	156
650	gp140	270	270
651	gp140	90	90
652	gp140	≥810	810
653	gp140	270	90
Geometric Mean Titer		270	207

15

<sup>1</sup>Reciprocal serum dilution at which HIV-induced syncytia of Sup T1 cells was inhibited by >90% compared to pre-immune serum. All prebleed sera were negative (titer <10).

**Table 5B**  
**Ability of Group M Consensus HIV-1 Envelope CON6 gp120 and gp140CF Proteins**  
**to Induce Antibodies that Neutralize HIV Primary Isolates**

HIV Isolate (Subtype)	CON6 gp120 Protein Guinea Pig No.					CON6 gp140CF Protein Guinea Pig No.					Controls		
	646	647	648	649	GMT	650	651	652	653	GMT	TriMab <sub>2</sub> †	CD4-IgG2	HIV+ Serum
SHIV 89.6P*(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	NT	NT	NT
SHIV SF162P3*(B)	<20	30	48	<20	<20	27	<20	<20	<20	<20	NT	0.2µg/ml	NT
BX08(B)	270	183	254	55	102	199	64	229	150	187	0.7µg/ml	NT	238#
6101(B)	<20	38	35	<20	<20	<20	90	72	73	39	1.1µg/ml	NT	NT
BG1168(B)	<20	<20	<20	<20	<20	40	<20	<20	25	<20	2.7µg/ml	NT	NT
0692(B)	31	32	34	<20	24	28	33	30	45	33	0.8µg/ml	NT	769
PAVO(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	2.9µg/ml	NT	NT
SF162(B)	2,146	308	110	282	379	206	5,502	15,098	174	1,313	NT	NT	>540
SS1196(B)	206	26	148	59	83	381	401	333	81	253	NT	NT	301#
BAL(B)	123	90	107	138	113	107	146	136	85	116	NT	NT	3307
92RW020(A)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	NT	NT	693
DU179(C)	<20	43	<20	24	<20	<20	<20	24	515	33	NT	0.8µg/ml	NT
DU368(C)	25	35	62	<20	27	<20	<20	<20	23	<20	NT	2.3µg/ml	NT
S021(C)	<20	<20	33	<20	<20	<20	<20	<20	<20	<20	NT	8.3µg/ml	NT
S080(C)	24	37	70	41	40	<20	<20	<20	52	<20	NT	3.4µg/ml	NT
93ZR001(D)	275	144	126	114	154	306	195	129	173	191	NT	NT	693
CM244(E)	35	43	64	ND	46	31	25	27	25	26	NT	NT	693

\*MT-2 Assay; All other HIV isolates were tested in the M7-luciferase assay.

HIV-1 isolates QH0692, SS1196, SF162, 6101, BX08, BG1168, BAL were assayed with post-injection 5 serum; other HIV-1 isolates were assayed with post-injection 4 serum. ND = not done.

HIV+ sera was either HIV-1+ human serum (LEH3) or an anti-gp120 guinea pig serum (#) with known neutralizing activity for HIV-1 isolate SS1196. GMT = geometric mean titer of four animals per group. Neutralizing titers reported are after subtraction of any background neutralization in prebleed sera.

†TriMab<sub>2</sub> = a mixture of human mabs 2F5, b12, 2G12.

CONCLUSIONS

The production of an artificial HIV-1 Group M consensus env genes (encoding sequences) (CON6 and  
5 Con-S) have been described that encodes a functional Env protein that is capable of utilizing the CCR5 co-receptor for mediating viral entry. Importantly, these Group M consensus envelope genes could induce T and B cell responses that recognized epitopes of  
10 subtype B and C HIV-1 primary isolates. In addition, Con-S induces antibodies that strongly neutralize Subtype-C and A HIV-1 strains (see Table 3).

The correlates of protection to HIV-1 are not  
15 conclusively known. Considerable data from animal models and studies in HIV-1-infected patients suggest the goal of HIV-1 vaccine development should be the induction of broadly-reactive CD4+ and CD8+ anti-HIV-1 T cell responses (Letvin et al, Annu.  
20 Rev. Immunol. 20:73-99 (2002)) and high levels of antibodies that neutralize HIV-1 primary isolates of multiple subtypes (Mascola et al, J. Virol. 73:4009-4018 (1999), Mascola et al, Nat. Med. 6:270-210 (2000)).

25 The high level of genetic variability of HIV-1 has made it difficult to design immunogens capable of inducing immune responses of sufficient breadth to be clinically useful. Epitope based vaccines for T and B cell responses (McMichael et al, Vaccine  
30 20:1918-1921 (2002), Sbai et al, Curr. Drug Targets Infect, Disord. 1:303-313 (2001), Haynes, Lancet

348:933-937 (1996)), constrained envelopes  
reflective of fusion intermediates (Fouts et al,  
Proc. Natl. Acad. Sci. USA 99:11842-22847 (2002)),  
as well as exposure of conserved high-order  
5 structures for induction of anti-HIV-1 neutralizing  
antibodies have been proposed to overcome HIV-1  
variability (Roben et al, J. Virol. 68:4821-4828  
(1994), Saphire et al, Science 293:1155-1159  
(2001)). However, with the ever-increasing  
10 diversity and rapid evolution of HIV-1, the virus is  
a rapidly moving complex target, and the extent of  
complexity of HIV-1 variation makes all of these  
approaches problematic. The current most common  
approach to HIV-1 immunogen design is to choose a  
15 wild-type field HIV-1 isolate that may or may not be  
from the region in which the vaccine is to be  
tested. Polyvalent envelope immunogens have been  
designed incorporating multiple envelope immunogens  
(Bartlett et al, AIDS 12:1291-1300 (1998), Cho et  
20 al, J. Virol. 75:2224-2234 (2001)).

The above-described study tests a new strategy  
for HIV-1 immunogen design by generating a group M  
consensus env gene (CON6) with decreased genetic  
distance between this candidate immunogen and wild-  
25 type field virus strains. The CON6 env gene was  
generated for all subtypes by choosing the most  
common amino acids at most positions (Gaschen et al,  
Science 296:2354-2360 (2002), Korber et al, Science  
288:1789-1796 (2000)). Since only the most common  
30 amino acids were used, the majority of antibody and  
T cell epitopes were well preserved. Importantly,



the genetic distances between the group M consensus env sequence and any subtype env sequences was about 15%, which is only half of that between wild-type subtypes (30%) (Gaschen et al, Science 296:2354-2360  
5 (2002)). This distance is approximately the same as that among viruses within the same subtype.

Further, the group M consensus env gene was also about 15% divergent from any recombinant viral env gene, as well, since CRFs do not increase the  
10 overall genetic divergence among subtypes.

Infectivity of CON6-Env pseudovirions was confirmed using a single-round infection system, although the infectivity was compromised, indicating the artificial envelope was not in an "optimal"  
15 functional conformation, but yet was able to mediate virus entry. That the CON6 envelope used CCR5 (R5) as its coreceptor is important, since majority of HIV-1 infected patients are initially infected with R5 viruses.

20 BIAcore analysis showed that both CON6 gp120 and gp140CF bound SCD4 and a number of mabs that bind to wild-type HIV-1 Env proteins. The expression of the CON6 gp120 and 140CF proteins that are similar antigenically to wild-type HIV-1  
25 envelopes is an important step in HIV-1 immunogen development. However, many wild-type envelope proteins express the epitopes to which potent neutralizing human mabs bind, yet when used as immunogens themselves, do not induce broadly  
30 neutralizing anti-HIV-1 antibodies of the specificity of the neutralizing human mabs.

The neutralizing antibody studies were encouraging in that both CON6 gp120, CON6 gp140CF and Con-S gp140CFI induced antibodies that neutralized select subtype B, C and D HIV-1 primary isolates, with Con-S gp140CFI inducing the most robust neutralization of non-subtype B primary HIV isolates. However, it is clear that the most difficult-to-neutralize primary isolates (PAVO, 6101, BG1168, 92RW020, CM244) were either only weakly or not neutralized by anti-CON6 gp120 or gp140 sera (Table 4b). Nonetheless, the Con-S envelope immunogenicity for induction of neutralizing antibodies is promising, given the breadth of responses generated with the Con-S subunit gp140CFI envelope protein for non-subtype B HIV isolates. Previous studies with poxvirus constructs expressing gp120 and gp160 have not generated high levels of neutralizing antibodies (Evans et al, J. Infect. Dis. 180:290-298 (1999), Polacino et al, J. Virol. 73:618-630 (1999), Ourmanov et al, J. Virol. 74:2960-2965 (2000), Pal et al, J. Virol 76:292-302 (2002), Excler and Plotkin, AIDS 11(Suppl A):S127-137 (1997). rVV expressing secreted CON6 gp120 and gp140 have been constructed and antibodies that neutralize HIV-1 primary isolates induced. An HIV neutralizing antibody immunogen can be a combination of Con-S gp140CFI, or subunit thereof, with immunogens that neutralize most subtype B isolates.

The structure of an oligomeric gp140 protein is critical when evaluating protein immunogenicity. In this regard, study of purified CON6 gp140CF proteins by fast performance liquid chromatography (FPLC) and analytical ultracentrifugation has demonstrated that the purified gp140 peak consists predominantly of trimers with a small component of dimers.

Thus, centralized envelopes such as CON6, Con-S or 2003 group M or subtype consensus or ancestral encoding sequences described herein, are attractive candidates for preparation of various potentially "enhanced" envelope immunogens including CD4-Env complexes, constrained envelope structures, and trimeric oligomeric forms. The ability of CON6-induced T and B cell responses to protect against HIV-1 infection and/or disease in SHIV challenge models will be studied in non-human primates.

The above study has demonstrated that artificial centralized HIV-1 genes such as group M consensus env gene (CON6) and Con-S can also induce T cell responses to T cell epitopes in wild-type subtype B and C Env proteins as well as to those on group M consensus Env proteins (Figure 5). While the DNA prime and rVV boost regimen with CON6 gp140CF immunogen clearly induced IFN- $\gamma$  producing T cells that recognized subtype B and C epitopes, further studies are needed to determine if centralized sequences such as are found in the CON6 envelope are significantly better at inducing cross-

clade T cell responses than wild-type HIV-1 genes  
(Ferrari et al, Proc. Natl. Acad. Sci. USA 94:1396-  
1401 (1997), Ferrari et al, AIDS Res. Hum.  
Retroviruses 16:1433-1443 (2000)). However, the  
5 fact that CON6 (and Con-S, env encoding sequence)  
prime and boosted splenocyte T cells recognized HIV-  
1 subtype B and C T cell epitopes is an important  
step in demonstration that CON6 (and Con-S) can  
induce T cell responses that might be clinically  
10 useful.

Three computer models (consensus, ancestor and  
center of the tree (COT)) have been proposed to  
generate centralized HIV-1 genes (Gaschen et al,  
Science 296:2354-2360 (2002), Gao et al, Science  
15 299:1517-1518 (2003), Nickle et al, Science  
299:1515-1517 (2003), Korber et al, Science  
288:1789-1796 (2000). They all tend to locate at  
the roots of the star-like phylogenetic trees for  
most HIV-1 sequences within or between subtypes. As  
20 experimental vaccines, they all can reduce the  
genetic distances between immunogens and field virus  
strains. However, consensus, ancestral and COT  
sequences each have advantages and disadvantages  
(Gaschen et al, Science 296:2354-2360 (2002), Gao et  
25 al, Science 299:1517-1518 (2003), Nickle et al,  
Science 299:1515-1517 (2003). Consensus and COT  
represent the sequences or epitopes in sampled  
current wild-type viruses and are less affected by  
outliers HIV-1 sequences, while ancestor represents  
30 ancestral sequences that can be significantly  
affected by outlier sequences. However, at present,

it is not known which centralized sequence can serve as the best immunogen to elicit broad immune responses against diverse HIV-1 strains, and studies are in progress to test these different strategies.

5 Taken together, the data have shown that the HIV-1 artificial CON6 and Con-S envelope can induce T cell responses to wild-type HIV-1 epitopes, and can induce antibodies that neutralize HIV-1 primary isolates, thus demonstrating the feasibility and  
10 promise of using artificial centralized HIV-1 sequences in HIV-1 vaccine design.

#### EXAMPLE 2

HIV-1 Subtype C Ancestral and Consensus Envelope  
15 Glycoproteins

#### EXPERIMENTAL DETAILS

HIV-1 subtype C ancestral and consensus env genes were obtained from the Los Alamos HIV Molecular Immunology Database (<http://hiv-web.lanl.gov/immunology>), codon-usage optimized for  
20 mammalian cell expression, and synthesized (Fig. 6). To ensure optimal expression, a Kozak sequence (GCCGCCGCC) was inserted immediately upstream of the initiation codon. In addition to the full-length  
25 genes, two truncated env genes were generated by introducing stop codons immediately after the gp41 membrane-spanning domain (IVNR) and the gp120/gp41 cleavage site (REKR), generating gp140 and gp120 form of the glycoproteins, respectively (Fig. 8).

Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. To determine if the ancestral and consensus subtype C envelopes were capable of mediating fusion and entry, *gp160* and *gp140* genes were co-transfected with an HIV-1/SG3Δenv provirus and the resulting pseudovirions tested for infectivity using the JC53-BL cell assay (Fig. 7). Co-receptor usage and envelope neutralization sensitivity were also determined with slight modifications of the JC53-BL assay. Codon-usage optimized and rev-dependent 96ZAM651 env genes were used as contemporary subtype C controls.

#### RESULTS

15

Codon-optimized subtype C ancestral and consensus envelope genes (*gp160*, *gp140*, *gp120*) express high levels of env glycoprotein in mammalian cells (Fig. 9).

20

Codon-optimized subtype C *gp160* and *gp140* glycoproteins are efficiently incorporated into virus particles. Western Blot analysis of sucrose-purified pseudovirions reveals ten-fold higher levels of virion incorporation of the codon-optimized envelopes compared to that of a rev-dependent contemporary envelope controls (Fig. 10A).

25

Virions pseudotyped with either the subtype C consensus *gp160* or *gp140* envelope were more infectious than pseudovirions containing the corresponding *gp160* and *gp140* ancestral envelopes.

30

Additionally, *gp160* envelopes were consistently more infectious than their respective *gp140* counterparts (Fig. 10B).

Both subtype C ancestral and consensus envelopes utilize CCR5 as a co-receptor to mediate virus entry (Fig. 11).

The infectivity of subtype C ancestral and consensus *gp160* containing pseudovirions was neutralized by plasma from subtype C infected patients. This suggests that these artificial envelopes possess a structure that is similar to that of native HIV-1 env glycoproteins and that common neutralization epitopes are conserved. No significant differences in neutralization potential were noted between subtype C ancestral and consensus env glycoproteins (*gp160*) (Fig. 12).

#### CONCLUSIONS

HIV-1 subtype C viruses are among the most prevalent circulating isolates, representing approximately fifty percent of new infections worldwide. Genetic diversity among globally circulating HIV-1 strains poses a challenge for vaccine design. Although HIV-1 Env protein is highly variable, it can induce both humoral and cellular immune responses in the infected host. By analyzing 70 HIV-1 complete subtype C env sequences, consensus and ancestral subtype C env genes have been generated. Both sequences are roughly equidistant from contemporary subtype C strains and thus

expected to induce better cross-protective immunity.  
A reconstructed ancestral or consensus sequence  
derived-immunogen minimizes the extent of genetic  
differences between the vaccine candidate and  
5 contemporary isolates. However, consensus and  
ancestral subtype C env genes differ by 5% amino  
acid sequences. Both consensus and ancestral  
sequences have been synthesized for analyses.  
Codon-optimized subtype C ancestral and consensus  
10 envelope genes have been constructed and the *in*  
*vitro* biological properties of the expressed  
glycoproteins determined. Synthetic subtype C  
consensus and ancestral env genes express  
glycoproteins that are similar in their structure,  
15 function and antigenicity to contemporary subtype C  
wild-type envelope glycoproteins.

### EXAMPLE 3

Codon-Usage Optimization of Consensus of Subtype C  
20 gag and nef Genes (C.con.gag and C.con.nef)

Subtype C viruses have become the most  
prevalent viruses among all subtypes of Group M  
viruses in the world. More than 50% of HIV-1  
25 infected people are currently carrying HIV-1 subtype  
C viruses. In addition, there is considerable  
intra-subtype C variability: different subtype C  
viruses can differ by as much as 10%, 6%, 17% and



16% of their Gag, Pol, Env and Nef proteins, respectively. Most importantly, the subtype C viruses from one country can vary as much as the viruses isolated from other parts of the world. The only exceptions are HIV-1 strains from India/China, Brazil and Ethiopia/Djibouti where subtype C appears to have been introduced more recently. Due to the high genetic variability of subtype C viruses even within a single country, an immunogen based on a single virus isolate may not elicit protective immunity against other isolates circulating in the same area.

Thus *gag* and *nef* gene sequences of subtype C viruses were gathered to generate consensus sequences for both genes by using a 50% consensus threshold. To avoid a potential bias toward founder viruses, only one sequence was used from India/China, Brazil and Ethiopia/Djibouti, respectively, to generate the subtype C consensus sequences (C.con.gag and C.con.nef). The codons of both C.con.gag and C.con.nef genes were optimized based on the codon usage of highly expressed human genes. The protein expression following transfection into 293T cells is shown in Figure 13. As can be seen, both consensus subtype C Gag and Nef proteins were expressed efficiently and recognized by Gag- and Nef-specific antibodies. The protein expression levels of both C.con.gag and C.con.nef genes are comparable to that of native subtype *env* gene (96ZM651).

EXAMPLE 4

Synthesis of a Full Length "Consensus of the  
Consensus env Gene with Consensus Variable Regions"  
(CON-S)

In the synthesized "consensus of the consensus"  
env gene (CON6), the variable regions were replaced  
with the corresponding regions from a contemporary  
subtype C virus (98CN006). A further con/con gene  
has been designed that also has consensus variable  
regions (CON-s). The codons of the Con-S env gene  
were optimized based on the codon usage of highly  
expressed human genes. (See Figs. 14A and 14B for  
amino acid sequences and nucleic acid sequences,  
respectfully.)

Paired oligonucleotides (80-mers) which overlap  
by 20 bp at their 3' ends and contain invariant  
sequences at their 5' and 3' ends, including the  
restriction enzyme sites EcoRI and BbsI as well as  
BsmBI and BamHI, respectively, were designed. BbsI  
and BamHI are Type II restriction enzymes that  
cleave outside of their recognition sequences. They  
have been positioned in the oligomers in such a way  
that they cleave the first four residues adjacent to  
the 18 bp invariant region, leaving 4 base 5'  
overhangs at the end of each fragment for the  
following ligation step. 26 paired oligomers were  
linked individually using PCR and primers  
complimentary to the 18 bp invariant sequences.

Each pair was cloned into pGEM-T (Promega) using the T/A cloning method and sequenced to confirm the absence of inadvertent mutations/deletions. pGEM-T subclones containing the proper inserts were then  
5 digested, run on a 1% agarose gel, and gel purified (Qiagen). Four individual 108-mers were ligated into pcDNA3.1 (Invitrogen) in a multi-fragment ligation reaction. The four-way ligations occurred among groups of fragments in a stepwise manner from  
10 the 5' to the 3' end of the gene. This process was repeated until the entire gene was reconstructed in the pcDNA3.1 vector.

A complete Con-S gene was constructed by ligating the codon usage optimized oligo pairs  
15 together. To confirm its open reading frame, an *in vitro* transcription and translation assay was performed. Protein products were labeled by <sup>35</sup>S-methionine during the translation step, separated on a 10% SDS-PAGE, and detected by radioautography.  
20 Expected size of the expressed Con-S gp160 was identified in 4 out of 7 clones (Fig. 14C).

CONs Env protein expression in the mammalian cells after transfected into 293T cells using a Western blot assay (Figure 15). The expression level  
25 of Con-S Env protein is very similar to what was observed from the previous CON6 env clone that contains the consensus conservative regions and variable loops from 98CN006 virus isolate.

The Env-pseudovirions was produced by  
30 cotransfecting Con-S env clone and env-deficient SG3

proviral clone into 293T cells. Two days after transfection, the pseudovirions were harvested and infected into JC53BL-13 cells. The infectious units (IU) were determined by counting the blue cells  
5 after staining with X-gal in three independent experiments. When compared with CON6 env clone, Con-S env clones produce similar number of IU in JC53BL-13 cells (Figure 16). The IU titers for both are about 3 log higher than the SG3 backbone clone  
10 control (No Env). However, the titers are also about 2 log lower than the positive control (the native HIV-1 env gene, NL4-3 or YU2). These data suggest that both consensus group M env clones are biologically functional. Their functionality,  
15 however, has been compromised. The functional consensus env genes indicate that these Env proteins fold correctly, preserve the basic conformation of the native Env proteins, and are able to be developed as universal Env immunogens.

20 It was next determined what coreceptor Con-S Env uses for its entry into JC53-BL cells. When treated with CXCR4 blocking agent AMD3100, the infectivity of NL4-3 Env-pseudovirions was blocked while the infectivity of YU2, Con-S or CON6 Env-pseudovirions was not inhibited. In contrast, when  
25 treated with CCR5 blocking agent TAK779, the infectivity of NL4-3 Env-pseudovirions was not affected, while the infectivity of YU2, Con-S or CON6 Env-pseudovirions was inhibited. When treated  
30 with both blocking agents, the infectivity of all pseudovirions was inhibited. Taken together, these

data show that the Con-S as well as CON6 envelope uses the CCR5 but not CXCR4 co-receptor for its entry into target cells.

It was next determined whether CON6 or Con-S Env proteins could be equally efficiently incorporated in to the pseudovirions. To be able precisely compare how much Env proteins were incorporated into the pseudovirions, each pseudovirions is loaded on SDS-PAGE at the same concentraion: 5µg total protein for cell lysate, 25ng p24 for cell culture supernatant, or 150ng p24 for purified virus stock (concentrated pseudovirions after super-speed centrifugation). There was no difference in amounts of Env proteins incorporated in CON6 or Con-S Env-pseudovirions in any preparations (cell lysate, cell culture supernatant or purified virus stock) (Figure 17).

#### EXAMPLE 5

Synthesis of a Consensus Subtype A Full Length *env*  
(A.con.env) Gene

Subtype A viruses are the second most prevalent HIV-1 in the African continent where over 70% of HIV-1 infections have been documented. Consensus *gag*, *env* and *nef* genes for subtype C viruses that are the most prevalent viruses in Africa and in the world were previously generated. Since genetic distances between subtype A and C viruses are as high as 30% in the *env* gene, the cross reactivity or protection between both subtypes will not be

optimal. Two group M consensus env genes for all subtypes were also generated. However, to target any particular subtype viruses, the subtype specific consensus genes will be more effective since the genetic distances between subtype consensus genes and field viruses from the same subtype will be smaller than that between group M consensus genes and these same viruses. Therefore, consensus genes need to be generated for development of subtype A specific immunogens. The codons of the A.con.env gene were optimized based on the codon usage of highly expressed human genes. (See Figs. 18A and 18B for amino acid and nucleic acid sequences, respectively.)

Each pair of the oligos has been amplified, cloned, ligated and sequenced. After the open reading frame of the A.con env gene was confirmed by an *in vitro* transcription and translation system, the A.con env gene was transfected into the 293T cells and the protein expression and specificity confirmed with the Western blot assay (Figure 18). It was then determined whether A.con envelope is biologically functional. It was co-transfected with the env-defective SG3 proviral clone into 293T cells. The pseudotyped viruses were harvested and used to infect JC53BL cells. Blue cells were detected in JC53-BL cells infected with the A.con Env-pseudovirions, suggesting that A.con Env protein is biologically functional (Table 6). However, the infectious titer of A.con Env-psuedovirions was about 7-fold lower than that of pseudovirions with

wild-type subtype C envelope (Table 6). Taken together, the biological function A.con Env proteins suggests that it folds correctly and may induce linear and conformational T and B cell epitopes if  
 5 used as an Env immunogen.

		JC53BL13 (IU/ul)		
		3/31/03	4/7/03	4/25/03
		non filtered supt.	0.22µm filtered	0.22µm filtered
A.con	+SG3	4	8.5	15.3
96ZM651	+SG3	87	133	104
SG3 backbone		0	0.07	0.03
Neg control		0	0.007	0

Table 6. Infectivity of pseudovirions with A.con env genes

#### EXAMPLE 6

10 Design of Full Length "Consensus of the Consensus gag, pol and nef Genes" (M.con.gag, M.con.pol and M.con.nef) and a Subtype C Consensus pol Gene (C.con.pol)

15 For the group M consensus genes, two different env genes were constructed, one with virus specific variable regions (CON6) and one with consensus variable regions (Con-S). However, analysis of T cell immune responses in immunized or vaccinated  
 20 animals and humans shows that the env gene normally is not a main target for T cell immune response

although it is the only gene that will induce neutralizing antibody. Instead, HIV-1 Gag, Pol and Nef proteins are found to be important for inducing potent T cell immune responses. To generate a repertoire of immunogens that can induce both broader humoral and cellular immune responses for all subtypes, it may be necessary to construct other group M consensus genes other than *env* gene alone. "Consensus of the consensus" *gag*, *pol* and *nef* genes (M.con.gag., M.con.pol and M.con.nef) have been designed. To generate a subtype consensus *pol* gene, the subtype C consensus *pol* gene (C.con.pol) was also designed. The codons of the M.con.gag., M.con.pol, M.con.nef and C.con.pol. genes were optimized based on the codon usage of highly expressed human genes. (See Fig. 19 for nucleic acid and amino acid sequences.)

#### EXAMPLE 7

##### Synthetic Subtype B Consensus *gag* and *env* Genes

#### 20 EXPERIMENTAL DETAILS

Subtype B consensus *gag* and *env* sequences were derived from 37 and 137 contemporary HIV-1 strains, respectively, codon-usage optimized for mammalian cell expression, and synthesized (Figs. 20A and 20B). To ensure optimal expression, a Kozak sequence (GCCGCCGCC) was inserted immediately upstream of the initiation codon. In addition to the full-length *env* gene, a truncated *env* gene was generated by introducing a stop codon immediately



after the gp41 membrane-spanning domain (IVNR) to create a *gp145* gene. Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. (Subtype B  
5 consensus Gag and Env sequences are set forth in Figs. 20C and 20D, respectively.)

To determine if the subtype B consensus envelopes were capable of mediating fusion and entry, *gp160* and *gp145* genes were co-transfected  
10 with an HIV-1/SG3Δenv provirus and the resulting pseudovirions were tested for infectivity using the JC53-BL cell assay. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also  
15 contain the reporter cassettes of luciferase and  $\beta$ -galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24-well plates, incubated at 37°C for 24  
20 hours and treated with DEAE-Dextran at 37°C for 30min. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-dextran, and allowed to incubate for 3 hours at 37°C after which an additional 500μL of cell media is added to each  
25 well. Following a final 48-hour incubation at 37°C, cells are fixed, stained using X-Gal, and overlaid with PBS for microscopic counting of blue foci. Counts for mock-infected wells, used to determine background, are subtracted from counts for the  
30 sample wells. Co-receptor usage and envelope

neutralization sensitivity were also determined with slight modifications of the JC53-BL assay.

To determine whether the subtype B consensus Gag protein was capable of producing virus-like particles (VLPs) that incorporated Env glycoproteins, 293T cells were co-transfected with subtype B consensus *gag* and *env* genes. 48-hours post-transfection, cell supernatants containing VLPs were collected, clarified in a tabletop centrifuge, filtered through a 0.2µm filter, and pellet through a 20% sucrose cushion. The VLP pellet was resuspended in PBS and transferred onto a 20-60% continuous sucrose gradient. Following overnight centrifugation at 100,000 x g, 0.5 ml fractions were collected and assayed for p24 content. The refractive index of each fraction was also measured. Fractions with the correct density for VLPs and containing the highest levels of p24 were pooled and pellet a final time. VLP-containing pellets were re-suspended in PBS and loaded on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from a subtype B HIV-1 infected individual.

## RESULTS

25

Codon-usage optimized, subtype B consensus envelope (*gp160*, *gp145*) and *gag* genes express high levels of glycoprotein in mammalian cells (Fig. 21).

Subtype B *gp160* and *gp145* glycoproteins are efficiently incorporated into virus particles.

30

Western Blot analysis of sucrose-purified pseudovirions suggests at least five-fold higher levels of consensus B envelope incorporation compared to incorporation of a rev-dependent contemporary envelope (Fig.23A). Virions pseudotyped with either the subtype B consensus gp160 or gp145 envelope are more infectious than pseudovirions containing a rev-dependent contemporary envelope (Fig. 23 B).

Subtype B consensus envelopes utilize CCR5 as the co-receptor to gain entry into CD4 bearing target cells (Fig. 22).

The infectivity of pseudovirions containing the subtype B consensus gp160 envelope was neutralized by plasma from HIV-1 subtype B infected patients (Fig. 24C) and neutralizing monoclonal antibodies (Fig. 24A). This suggests that the subtype B synthetic consensus B envelopes is similar to native HIV-1 Env glycoproteins in its overall structure and that common neutralization epitopes remain intact. Figs. 24B and 24D show neutralization profiles of a subtype B control envelope (NL4.3 Env).

Subtype B consensus Gag proteins are able to bud from the cell membrane and form virus-like particles (Fig. 25A). Co-transfection of the codon-optimized subtype B consensus gag and gp160 genes produces VLPs with incorporated envelope (Fig. 25B).

CONCLUSIONS

The synthetic subtype B consensus *env* and *gag* genes express viral proteins that are similar in their structure, function and antigenicity to contemporary subtype B Env and Gag proteins. It is contemplated that immunogens based on subtype B consensus genes will elicit CTL and neutralizing immune responses that are protective against a broad set of HIV-1 isolates.

10

\* \* \*

All documents and other information sources cited above are hereby incorporated in their entirety by reference. Also incorporated by reference is Liao et al, J. Virol. 78:5270 (2004)).

WHAT IS CLAIMED IS:

1. An isolated protein comprising the sequence of amino acids set forth in Fig. 1A.
2. A nucleic acid comprising a nucleotide sequence encoding CON6 HIV gp160 protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.
3. The nucleic acid according to claim 2 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 1D.
4. A nucleic acid comprising a nucleotide sequence encoding subtype C ancestral HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.
5. The nucleic acid according to claim 4 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 6A.
6. A nucleic acid comprising a nucleotide sequence encoding subtype C consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.
7. The nucleic acid according to claim 6 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 6B.
8. An isolated protein comprising the sequence of amino acids set forth in Fig. 6C or 6D.

9. A nucleic acid comprising a nucleotide sequence encoding a subtype C consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

10. The nucleic acid according to claim 9 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 13E.

11. A nucleic acid comprising a nucleotide sequence encoding a subtype C consensus HIV nef protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

12. The nucleic acid according to claim 11 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 13F.

13. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

14. The nucleic acid according to claim 13 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 14B.

15. A nucleic acid comprising a nucleotide sequence encoding subtype A consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

16. The nucleic acid according to claim 15 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 18B.

17. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

18. The nucleic acid according to claim 17 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19A .

19. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV pol protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

20. The nucleic acid according to claim 19 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19B.

21. A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV nef protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

22. The nucleic acid according to claim 21 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19C.

23. A nucleic acid comprising a nucleotide sequence encoding subtype C consensus HIV pol

protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

24. The nucleic acid according to claim 23 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19D.

25. A nucleic acid comprising a nucleotide sequence encoding subtype B consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

26. The nucleic acid according to claim 25 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 20A.

27. A nucleic acid comprising a nucleotide sequence encoding subtype B consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

28. The nucleic acid according to claim 27 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 20B.

29. An isolated protein comprising the sequence of amino acids set forth in Fig. 20C or 20D.

30. An isolated protein comprising the sequence of amino acids set forth in Fig. 26A .



31. A nucleic acid comprising a nucleotide sequence that encodes the protein according to claim 30.

32. The nucleic acid according to claim 31 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 26B.

33. An isolated protein comprising the sequence of amino acids set forth in Fig. 28B.

34. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 33.

35. The nucleic acid sequence according to claim 34 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 28C.

36. An isolated protein comprising the sequence of amino acids set forth in Fig. 29B.

37. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 36.

38. The nucleic acid sequence according to claim 37 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 29C.

39. An isolated protein comprising the sequence of amino acids set forth in Fig. 30B.

40. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 39.

41. The nucleic acid sequence according to claim 40 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 30C.

42. An isolated protein comprising the sequence of amino acids set forth in Fig. 31B.

43. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 42.

44. The nucleic acid sequence according to claim 43 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 31C.

45. An isolated protein comprising the sequence of amino acids set forth in Fig. 32B.

46. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 45.

47. The nucleic acid sequence according to claim 46 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 32C.

48. An isolated protein comprising the sequence of amino acids set forth in Fig. 33B.

49. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 48.

50. The nucleic acid sequence according to claim 49 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 33C.

51. An isolated protein comprising the sequence of amino acids set forth in Fig. 34B.

52. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 51.

53. The nucleic acid sequence according to claim 52 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 34C.

54. An isolated protein comprising the sequence of amino acids set forth in Fig. 35B.

55. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 54.

56. The nucleic acid sequence according to claim 55 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 35C.

57. An isolated protein comprising the sequence of amino acids set forth in Fig. 36B.

58. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 57.

59. The nucleic acid sequence according to claim 58 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 36C.

60. An isolated protein comprising the sequence of amino acids set forth in Fig. 37B.

61. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 60.

62. The nucleic acid sequence according to claim 61 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 37C.

63. An isolated protein comprising the sequence of amino acids set forth in Fig. 38B.

64. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 63.

65. The nucleic acid sequence according to claim 64 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 38C.

66. An isolated protein comprising a CF or CFI form of the amino acid sequence set forth in any one of Figs. 39A-127A.

67. A nucleic acid comprising the nucleotide sequence set forth in Fig. 39B.

68. A nucleic acid comprising the nucleotide sequence set forth in Fig. 40B.

69. A nucleic acid comprising the nucleotide sequence set forth in Fig. 41B.

70. A nucleic acid comprising the nucleotide sequence set forth in Fig. 42B.

71. A nucleic acid comprising the nucleotide sequence set forth in Fig. 43B.

72. A nucleic acid comprising the nucleotide sequence set forth in Fig. 44B.

73. A nucleic acid comprising the nucleotide sequence set forth in Fig. 45B.

74. A nucleic acid comprising the nucleotide sequence set forth in Fig. 46B.

75. A nucleic acid comprising the nucleotide sequence set forth in Fig. 47B.

76. A nucleic acid comprising the nucleotide sequence set forth in Fig. 48B.

77. A nucleic acid comprising the nucleotide sequence set forth in Fig. 49B.

78. A nucleic acid comprising the nucleotide sequence set forth in Fig. 50B.

79. A nucleic acid comprising the nucleotide sequence set forth in Fig. 51B.

80. A nucleic acid comprising the nucleotide sequence set forth in Fig. 52B.

81. A nucleic acid comprising the nucleotide sequence set forth in Fig. 53B.

82. A nucleic acid comprising the nucleotide sequence set forth in Fig. 54B.

83. A nucleic acid comprising the nucleotide sequence set forth in Fig. 55B.

84. A nucleic acid comprising the nucleotide sequence set forth in Fig. 56B.

85. A nucleic acid comprising the nucleotide sequence set forth in Fig. 57B.
86. A nucleic acid comprising the nucleotide sequence set forth in Fig. 58B.
87. A nucleic acid comprising the nucleotide sequence set forth in Fig. 59B.
88. A nucleic acid comprising the nucleotide sequence set forth in Fig. 60B.
89. A nucleic acid comprising the nucleotide sequence set forth in Fig. 61B.
90. A nucleic acid comprising the nucleotide sequence set forth in Fig. 62B.
91. A nucleic acid comprising the nucleotide sequence set forth in any one of Figs. 63B-84B, 65D, 67D and 68D.
92. A nucleic acid comprising the nucleotide sequence set forth in any one of Figs. 85B-106B, 88D, 90D and 92D.
93. A nucleic acid comprising the nucleotide sequence set forth in any one of Figs. 107B-127B, 109D, 111D and 112D.
94. A vector comprising the nucleic acid according to any one of claims 2-7, 9-28, 31, 32, 34, 35, 37, 38, 40, 41, 43, 44, 46, 47, 49, 50, 52, 53, 55, 56, 58, 59, 61, 62, 64, 65 and 67-93.

95. A composition comprising at least one protein or nucleic acid according to any one of claims 1-93 and a carrier.

96. A method of inducing an immune response in a mammal comprising administering to said mammal an amount of at least one protein and/or nucleic acid according to any one of claims 1-93 sufficient to effect said induction.

THIS PAGE BLANK (USPTO)



1/178

MRVMGIQRNCQHLWRWGTMLMLMICSAAENLWTVVYGVVPVWKEANTTLFCASDAKAYDTEVHNVWAT  
 HACVPTDPNPQEI VLE NV TEN EN FN MW KNN MV EQ MH E D I I S L W D Q S L K P C V K L T P L C V T L N C T N V R N V S S N G  
 V1  
 TETDNEEIKNC SFNITTELDRKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYRLINCN TSAITQACP  
 V2  
 KVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHG IKPVVSTQLLINGSLAEEEEIIIRSEN  
 V3  
 ITNNAKTIIVQLNESVEINCTRPNNNTRKRSIHIGPGQAFYATGEIIGDIRQAHCNISRTKWNKTLQQVAK  
 V4  
 KLREHENNKTIIFKPSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWWMFNCTYMFNGTKDNSETITLPCR  
 V5  
 IKQIINMWQGVGOAMYAPPIEGKITCKSNITGLLLTRDGGNNSNKNKTETETFRPGGDMRDNRSELYKYK  
 VVKIEPLGVAPTAKRRVVEREKRAVGIGAVELGELGAAGSTMGAASITLTVQARQLLSGIVQQQSLLR  
 AIEAQQLLQLTVWGIKQLQARVLAVERYLKDQQLLGWC SGKLICTTNVPWNSSWSNKSQDEIWDNMT  
 WMEWEREISNYTDIIYRLIEESQNQQEKNEQELLALDKWASLWNWFEDITNWLWYIKIFIMIVGGLIGLRI  
 VFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDPRPEGIEEGEGEQGRDRSIRLVNGFLALAWDDLRSLCLFS  
 YHRLRDFILIAARTVELLGRRLRGLQKGWEALKYLG NLLQYWGQELKN SAI SLLD TTAIAVAEGTDRVI  
 EIVQACRAILNIPRRIRQGLERALL

Fig. 1A

2/178

Fig. 1B

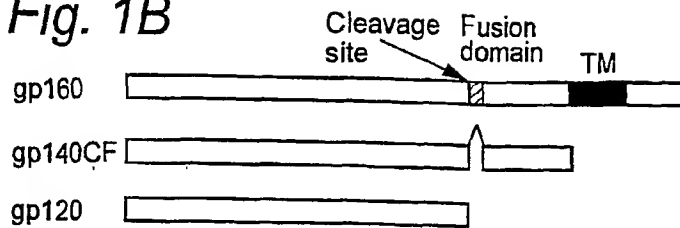


Fig. 1C

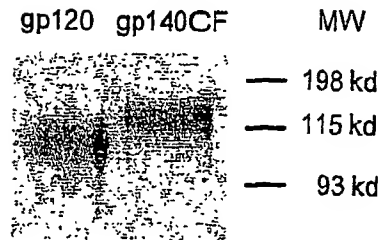


Fig. 1D

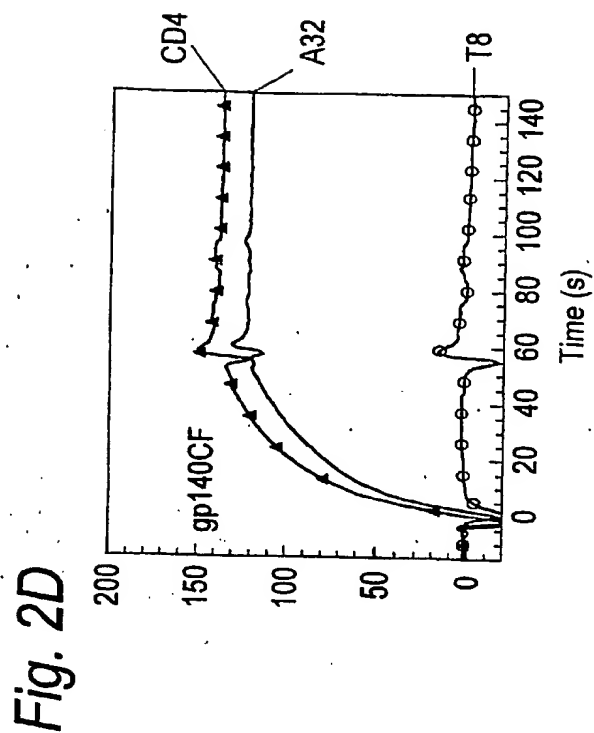
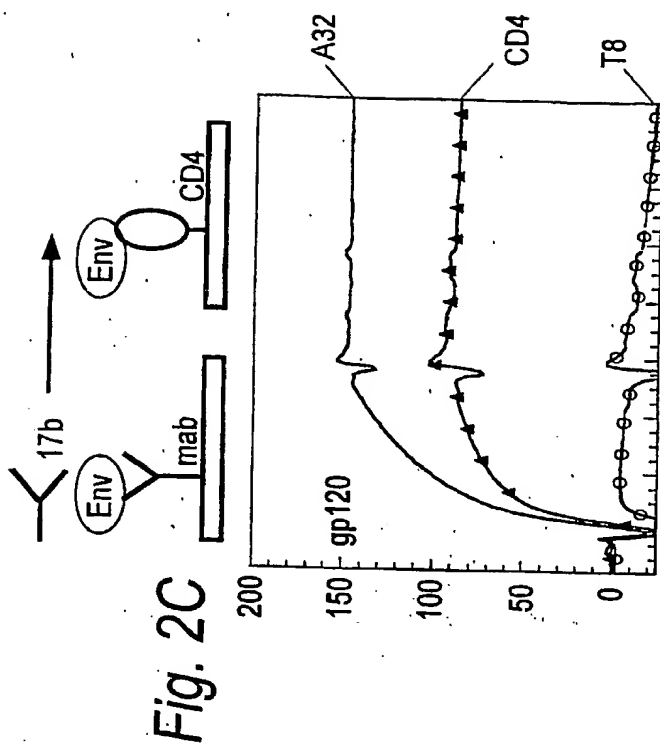
CON6.env (group M env consensus. This one contain five variable regions in env gene from 98CN006 virus, not in the public domain yet)

```

GCCACCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCATGATC
CTGGGCATGCTGATGATCTGCTCCGCGCCGAGAACCTGTGGGTGACCGTGTACTACGGC
GTGCCCCGTGTGGAAGGAGGCCAACACCCACCTGTTCTGCGCCTCCGACGCCAAGGCCTAC
GACACCGAGGTGCACAACGTGTGGGCCACCCACGCTGCGTGCCACCGACCCCAACCC
CAGGAGATCGTGCTGGAGAACGTGACCGGAGAACTTCAACATGTGGAAGAACAACATGGTG
GAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAG
CTGACCCCCCTGTGCGTGACCCCTGAAGTGCACCAACGTGCGCAACGTGTCTCCAACGGC
ACCGAGACCGACAACGAGGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGCTGCGC
GACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCTGGACGTGGTGCCCATCGACGAC
AAGAACTCCTCCGAGATCTCCGGCAAGAACTCCTCCGAGTACTACCGCTGATCAACTGC
AACACCTCCGCCATCACCCAGGCCCTGCCCAAGGTGTCTTTCGAGCCCATCCCCATCCAC
TACTGCGCCCCCGCCGGCTTCCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCCTGCAAGAACGTGTCCACCGTGACGTGCACCCACGGCATCAAGCCCGTGGTGTCC
ACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAAC
ATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGC
ACCCGCCCCAACAACAACACCCGCAAGTCCATCCACATCGGCCCCGGCCAGGCCCTTCTAC
GCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCCGCACCAAG
TGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACC
ATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGC
GGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATGTTCAAC
GGCACCTACATGTTCAACGGCACCAAGGACAACCTCCGAGACCATCACCTGCCCTGCCG
ATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCATC
GAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCCTGCTGCTGACCCGCGACGGCGGC
AACAACTCCAACAAGAACAAGACCGAGACCTTCCGCCCCGGCGGGCGGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCC
CCCACCAAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCC
GTGTTCTCTGGGCTTCTTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCTG
ACCGTGACGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGC
GCCATCGAGGCCCGAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAG
GCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGGACCAAGCAGCTGCTGGGCATCTGGGGC
TGCTCCGGCAAGCTGATCTGCACCACCAACGTGCCCTGGAACCTCCTCCTGGTCCAACAAG
TCCCAGGACGAGATC TGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCTCCAAC
TACACCGACATCATCTACCGCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAG
CAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACCTGGTTCGACATCACCAAC
TGGCTGTGGTATACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATC
GTGTTCCGCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTTC
CAGACCCTGATCCCCAACCCCGCGCGCCGACCGCCCGAGGGCATCGAGGAGGAGGGC
GGCGAGCAGGGCCGCGACCGCTCCATCCGCTGGTGAACGGCTTCTTGCCCTGGCCTGG
GACGACCTGCGCTCCCTGTGCTGTCTCTACCACCGCCTGCGCGACTTCATCCTGATC
GCCGCCCGCACCGTGGAGCTGCTGGGCCGCGCTCCCTGCGCGGCTGCAGAAGGGCTGG
GAGGCCCTGAAGTACCTGGGCAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAATCC
GCCATCTCCCTGCTGGACACCACCGCCATCGCCCTGGCCGAGGGCACCGACCGGTGATC
GAGATCGTGCAGCGCGCCTGCGCGCCATCCTGAACATCCCCCGCCCATCCGCCAGGGC
CTGGAGCGCGCCCTGCTGTAA

```

3/178



4/178

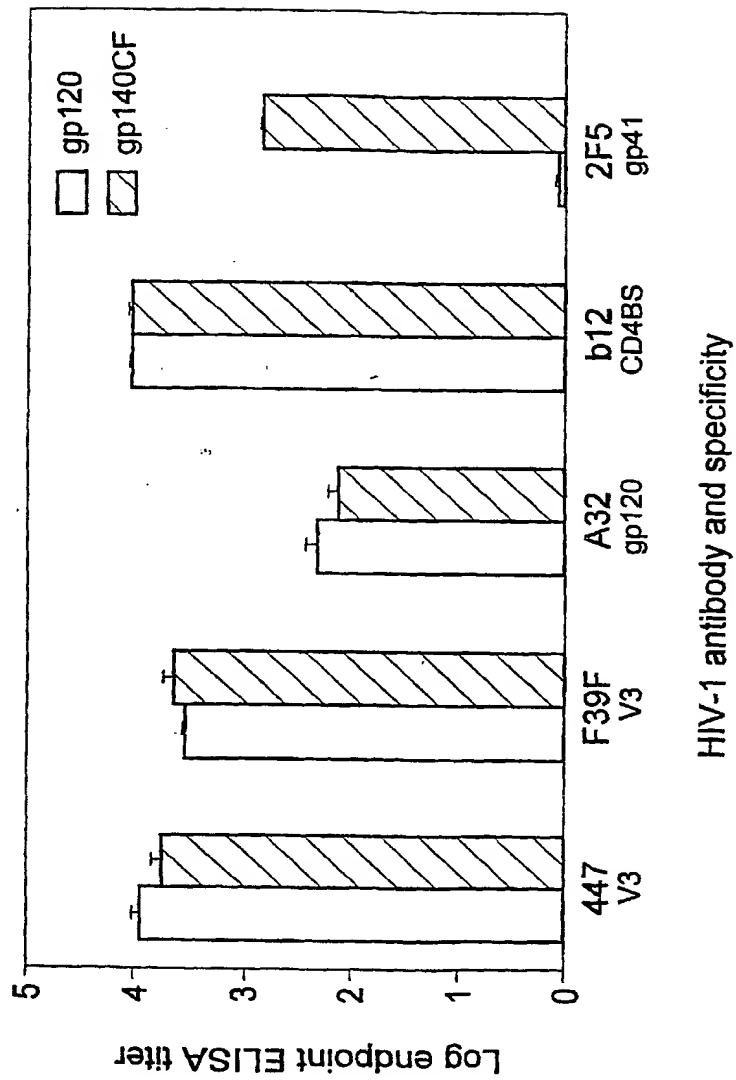
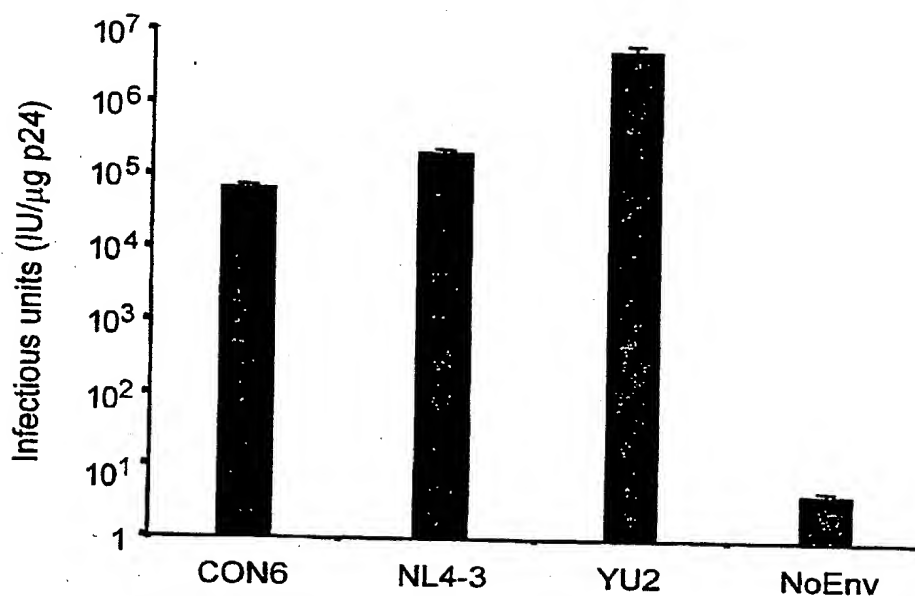
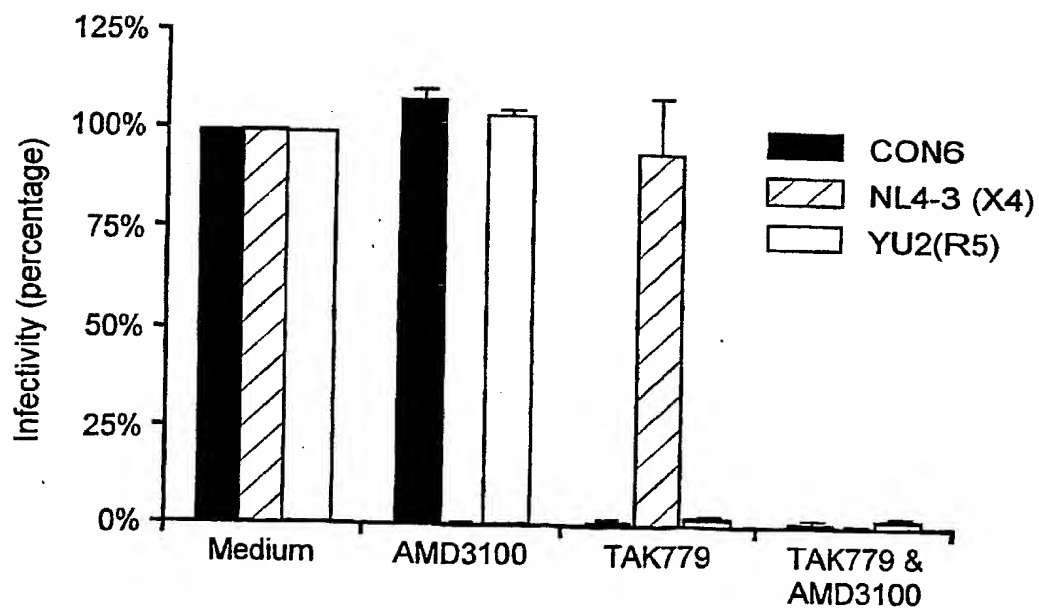


Fig. 2E

5/178

*Fig. 3A**Fig. 3B*

6/178

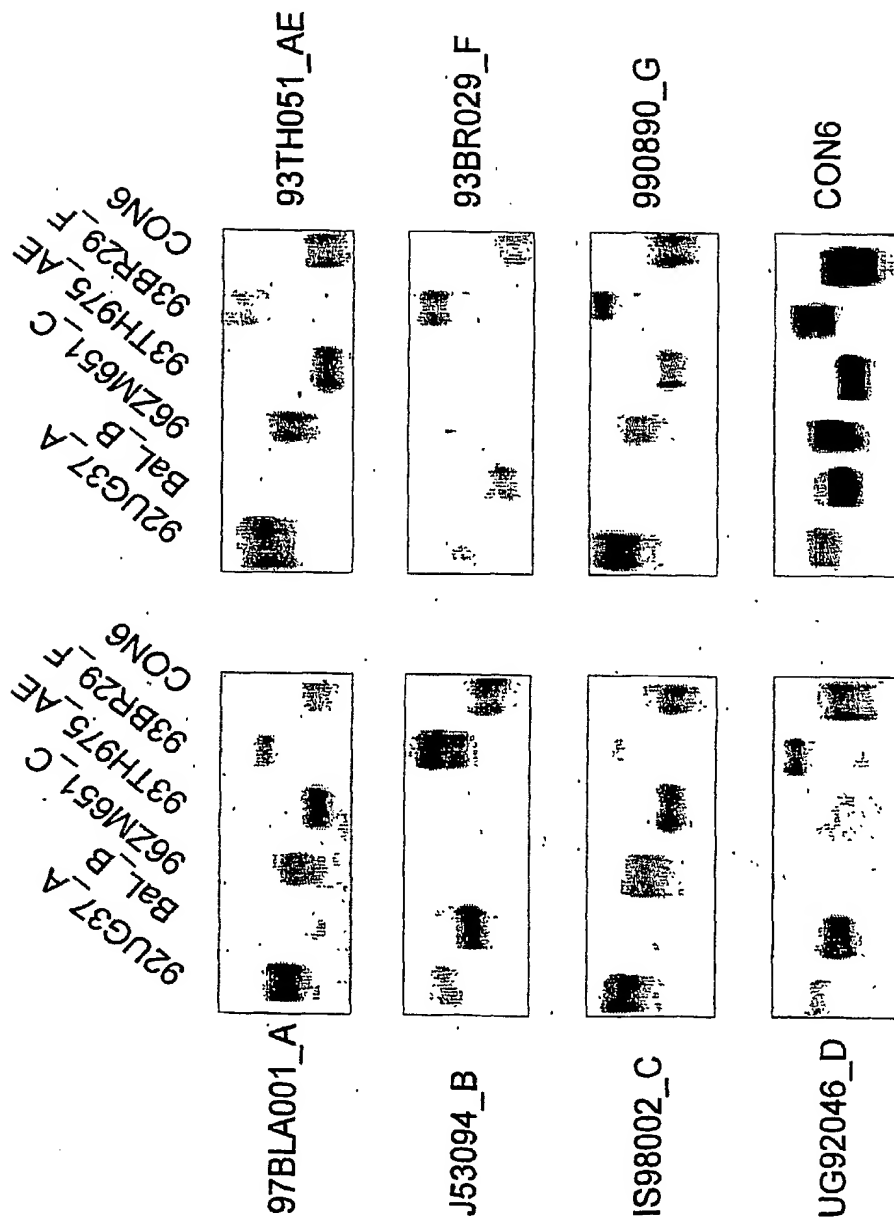
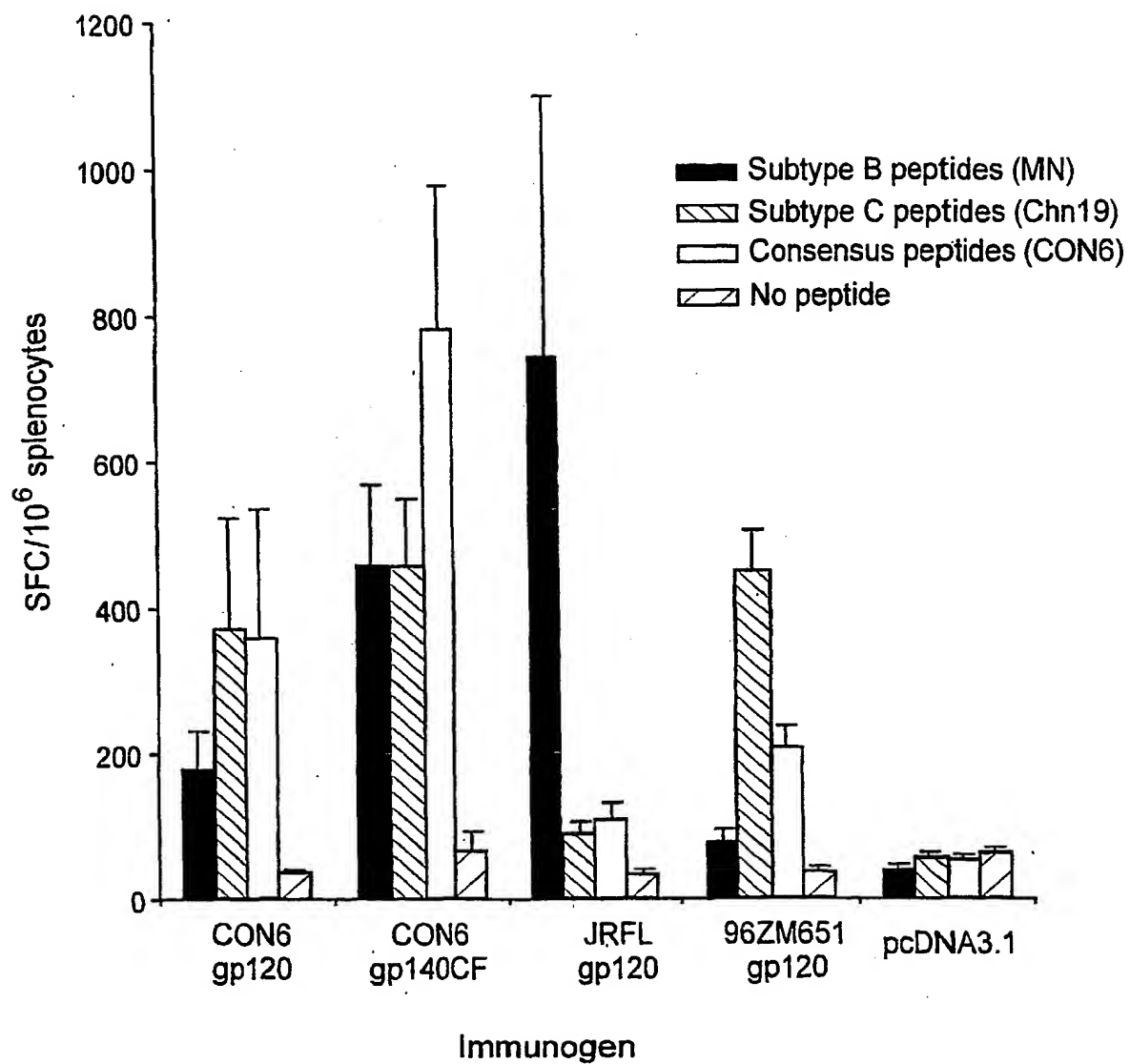


Fig. 4

7/178

*Fig. 5*

**Fig. 6A**

[illegible]



9/178

*Fig. 6B*

C.con.env (subtype C consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT  
 CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTGGTGGGCA  
 ACCTGTGGGTGACCGTGTAACGCGCGTGCCCGTGTTGGAAGGAGGCCAAG  
 ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGG AGGTGCA  
 CAACGTGTGGGGCACCACGCCTGCGTGCCACCGACCCCAACCCCAAGG  
 AGATGGTGTCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC  
 ATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCT  
 GAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCCGCA  
 ACGTGACCAACGCCACCAACAACACCTACAACGAGGAGATCAAG AACTGC  
 TCCTTCAACATCACCACCGAGCTGCGCGACAAGAAGAAGGTGTACGC  
 CCTGTTCTACCGCCTGGACATCGTGCCCTGAACGAGAACTCCTCCGAGT  
 ACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCAAG  
 GTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGC  
 CATCCTGAAGTGCAACAACAAGACCTTCAACGGCACCGGCCCTG CAACA  
 ACGTGTCCACCGTGCACTGCACCCACGGCATCAAGCCCGTGGTGTCCACC  
 CAGCTGTCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTC  
 CGAGAACCTGACCAACAACGCCAAGACCATCATCGTGACCTGAACGAGT  
 CCGTGGAGATCGTGTGCACCCGCCCAACAACAACACCCGCAAGTCCATC  
 CGCATCGGCCCCGGCCAGACCTTCTACGCCACCGCGGACATCATCG GCGA  
 CATCCGCCAGGCCCACTGCAACATCTCCGAGGACAAGTGGAACAAGACCC  
 TGCAGCGCGTGTCCAAGAAGCTGAAGGAGCACTTCCCCAACAAGACCATC  
 AAGTTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCAACCCACTCCTT  
 CAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACT  
 CCACCTACAACAACAACACCAACTCCAACTCCACCATCACCTGCCC TGC  
 CGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGTA  
 CGCCCCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCGGCC  
 TGCTGCTGACCCGCGACGGCGGCAAGAAGAACACCACCGAGATCTTCCGC  
 CCCGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTA  
 CAAGGTGGTGGAGATCAAGCCCTGGGCGTGGCCCCCAACAAGGCCAA GC  
 GCCGCTGGTGGAGCGCGAGAAGCGCGCGTGGGCATCGGCGCGTGTTC  
 CTGGGCTTCTGGGCGCGCGGCTCCACCATGGGCGCGCGCTCCATCAC  
 CCTGACCGTGCGAGGCCCGCCAGCTGCTGTCCGGCATCGTGACGACGAGT  
 CCAACCTGTGCGCGCCATCGAGGCCAGCAGCACATGCTGCAGCTGACC  
 GTGTGGGGCATCAAGCAGCTGCAGACCCGCGTGTGGCCATCGAGCGCTA  
 CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGCTGCTCCGGCAAGCTGA  
 TCTGCACCAACCGCGTGGCCTGGAACCTCCTCCTGGTCCAACAAGTCCCAG  
 GAGGACATCTGGGACAACATGACCTGGATGCAAGTGGGACCGCGAGATCTC  
 CAACTACACCGACACCATCTACCGCTGCTGGAGGACTCCCAGAACCAGC  
 AGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACTCCTGGAAGAACCCTG  
 TGAAGTGGTTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT  
 CATGATCGTGGGCGGCTGATCGGCCCTGCGCATCATCTTCGCCGTGCTGT  
 CCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACC  
 CTGACCCCAACCCCGCGGCCCGACCGCTGGGCCGATCGAGGAGGA  
 GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCTGGTGTCCGGCTTCC  
 TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCAC  
 CGCCTGCGCGACTTCATCCTGGTGGCCGCCCGCGCGCTGGAGCTGCTGGG  
 CCGCTCCTCCCTGCGCGGCTGCGCGCGCTGGGAGGCCCTGAAGTACC  
 TGGGCTCCCTGGTGCAGTACTGGGCTGGAGCTGAAGAAGTCCGCCATC  
 TCCCTGTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCAT  
 CATCGAGCTGATCCAGCGCATCTGCCGCGCCATCCGCAACATCCCCCGCC  
 GCATCCGCGACGCTGGAGGCGCGCTTCAGTAA

10/178

C.anc.env (subtype C ancestral env)

MRVMGILRNCQQWIIWGIIGFWMLMI CSVVGNLWTVYGVVPVWKEAKTTLCASDAKAYEREVHNVMAT  
HACVPTDPNPQEMVLENVTENFNWKNMVDQMHEDI ISLWDQSLKPCVKLTPLCVTLNCTNVTNATNNT  
YNGEMKNCSEFNI TELRDKKKKKEVALFYRLDIVPLN ENSSEYRLINCNTSAITQACPKVSFDPPIPIHYCA  
PAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEEIIIRSENLTDNAKTIIVQLN  
ESVEIVCTRPNNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISEDKNKTLQQAELKGFHFNKTIIF  
EPSSGGDLEITTHSFNCRGEFFYCNSTKLFNSTYNNNTNSNTITLPCRKQIINMWQGVGQAMYPPIA  
GNITCKSNITGLLLTRDGGKENTTETFRPGGDMRDNRSELYKYKVVEIKPLGVAPTEAKRRVVEREKR  
AVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHMLQLTWGIKQLQARVL  
AMERYLKDQQLLGIWGCCKLICTTAVPWNSSWSNKSLLDDIWDNMTWMEWDREISNYTDTIYRLLLEESQN  
QQEKNEQDLLALDSWENLWNNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLT  
PNPRGPDRLRIIEEGEGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLGRSSLR  
GLQRGWEALKYLGSLVQYWGWQELKKSALSLDDTIAIAVAEGTDRIIEVVQACRAILNIPRRIRQGFEEA  
LL

Fig. 6C

C.con.env (subtype C consensus env)

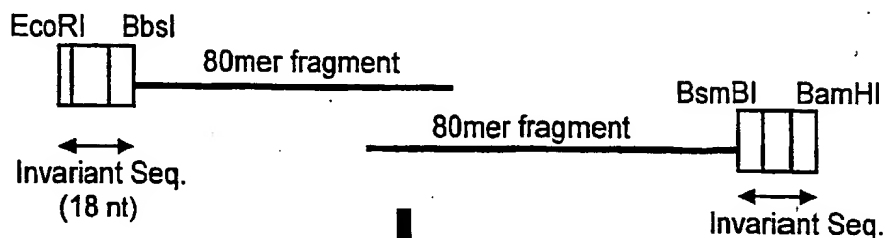
MRVMGILRNCQQWIIWGIIGFWMLMI CSVVGNLWTVYGVVPVWKEA KTTLCASDAKAYEREVHNVMAT  
HACVPTDPNPQEMVLENVTENFNWKNMVDQMHEDI ISLWDQSLKPCVKLTPLCVTLNCRNVTNATNNT  
YNEEIKNCSEFNI TELRDKKKKVALFYRLDIVPLNENSSSEYRLINCNTSAITQACPKVSFDPPIPIHYCA  
PAGYAILKCNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLNGSLAEEIIIRSENLTNAKTIIVHLN  
ESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEDKNKTLQRVSKKLKEHFPNKTIF  
EPSSGGDLEITTHSFNCRGEFFYCNSTKLFNSTYNNNTNSNTITLPCRKQIINMWQEVGRAMYPPIA  
GNITCKSNITGLLLTRDGGKNTTEIFRPGGDMRDNRSELYKYKVVEIKPLGVAPTKAKRRVVEREKR  
AVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHMLQLTWGI KQLQTRVL  
AIERYLKDQQLLGIWGCCKLICTTAVPWNSSWSNKSQEDIDWNMTWQWDREISNYTDTIYRLLLEDSQN  
QQEKNEKDLLALDSWKNLWNNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLT  
PNPRGPDRLRIIEEGEGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILVAARAVELLGRSSLR  
GLQRGWEALKYLGSLVQYWGWLELKKSAISLLDTIAIAVAEGTDRIIELIQICRAIRNIPRRIRQGFEEA  
LQ

Fig. 6D

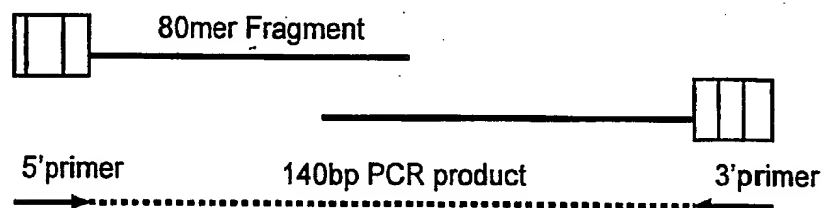
11/178

*Fig. 6E*

Synthesize entire gene in 80-mer fragments overlapping by 20 residues at the 3' end with invariant sequences at the 5' end.

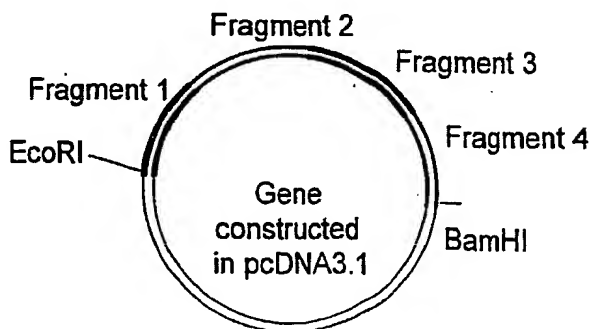


Paired 80mer oligos are connected via PCR in a stepwise manner from 5' to 3' using primers complimentary to the invariant seq.



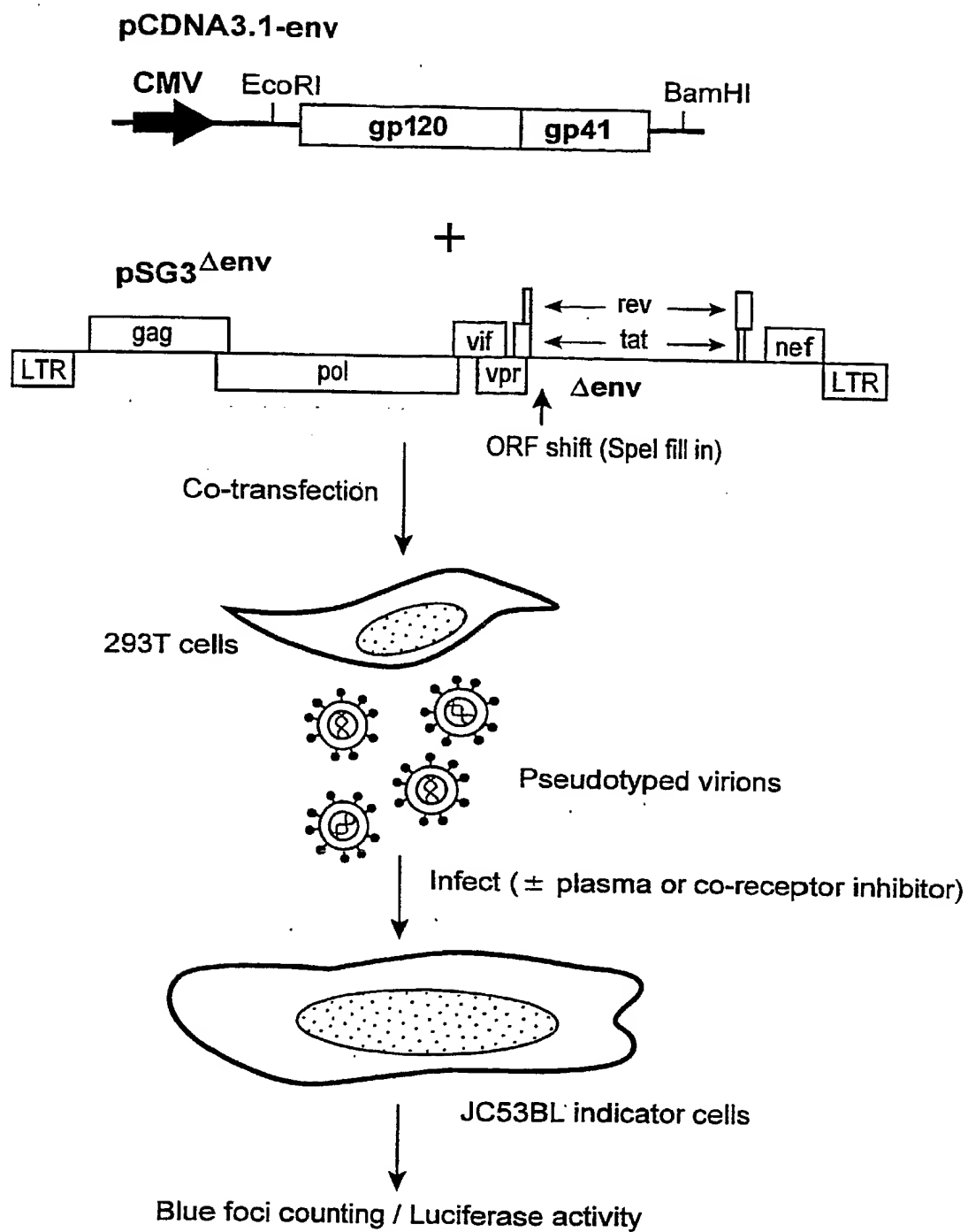
108bp PCR fragments cloned into pGEM-T and sequenced. Clones with the proper sequence will be cut with 2 restriction enzymes. 4 fragments will be ligated together with pcDNA3.1 in a stepwise manner from the 5' to 3' end of gene

Fragments to be ligated with pcDNA3.1 (1-4 are in order from 5' to 3')	Restriction Enzymes Used to Cleave Fragment
Fragment 1	EcoRI/BsmBI
Fragment 2	BbsI/BsmBI
Fragment 3	BbsI/BsmBI
Fragment 4	BbsI/BamHI
pcDNA3.1	EcoRI/BamHI



Ligations will be repeated stepwise 5' to 3' until the entire gene has been cloned into pcDNA3.1

12/178

*Fig. 7*



14/178

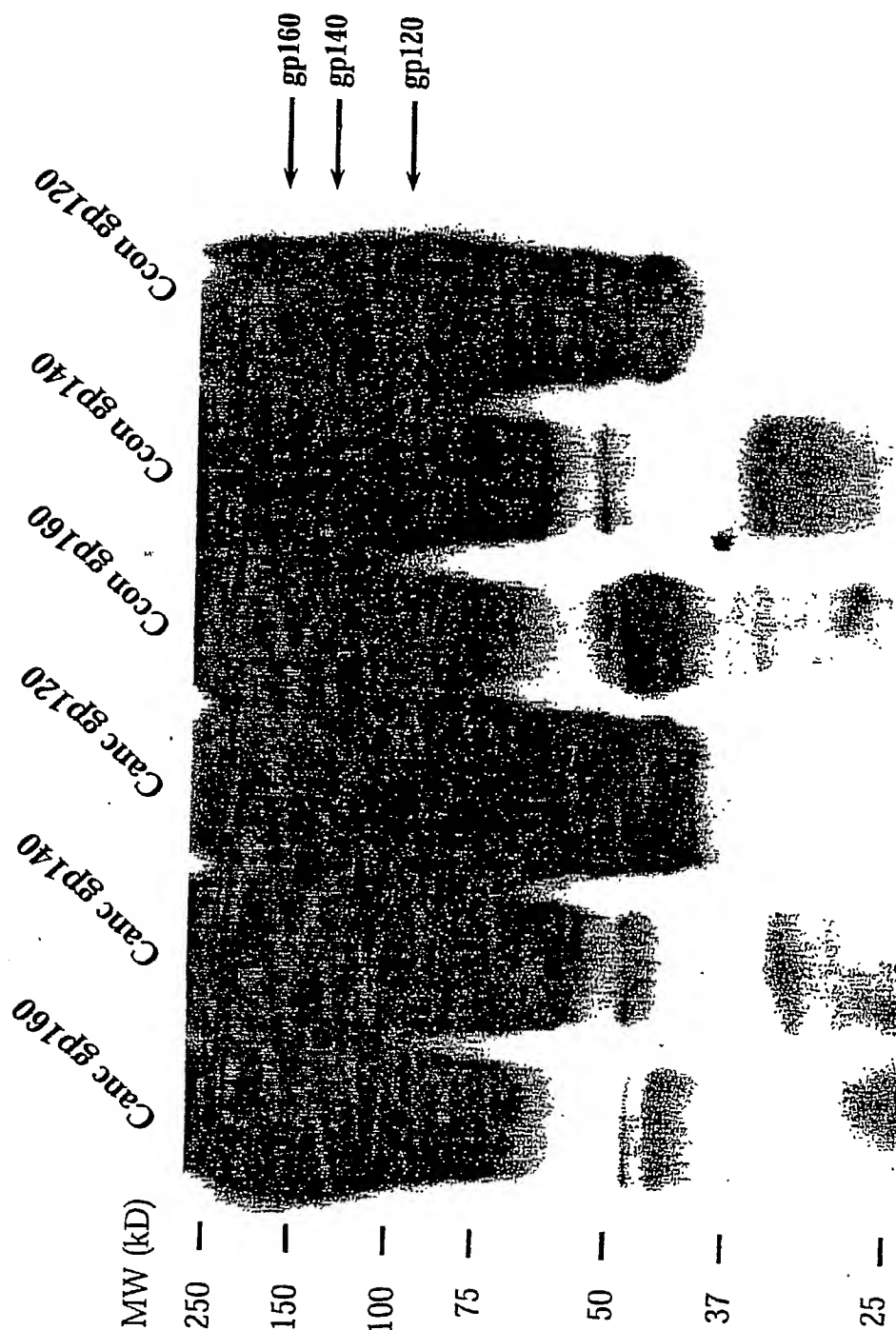
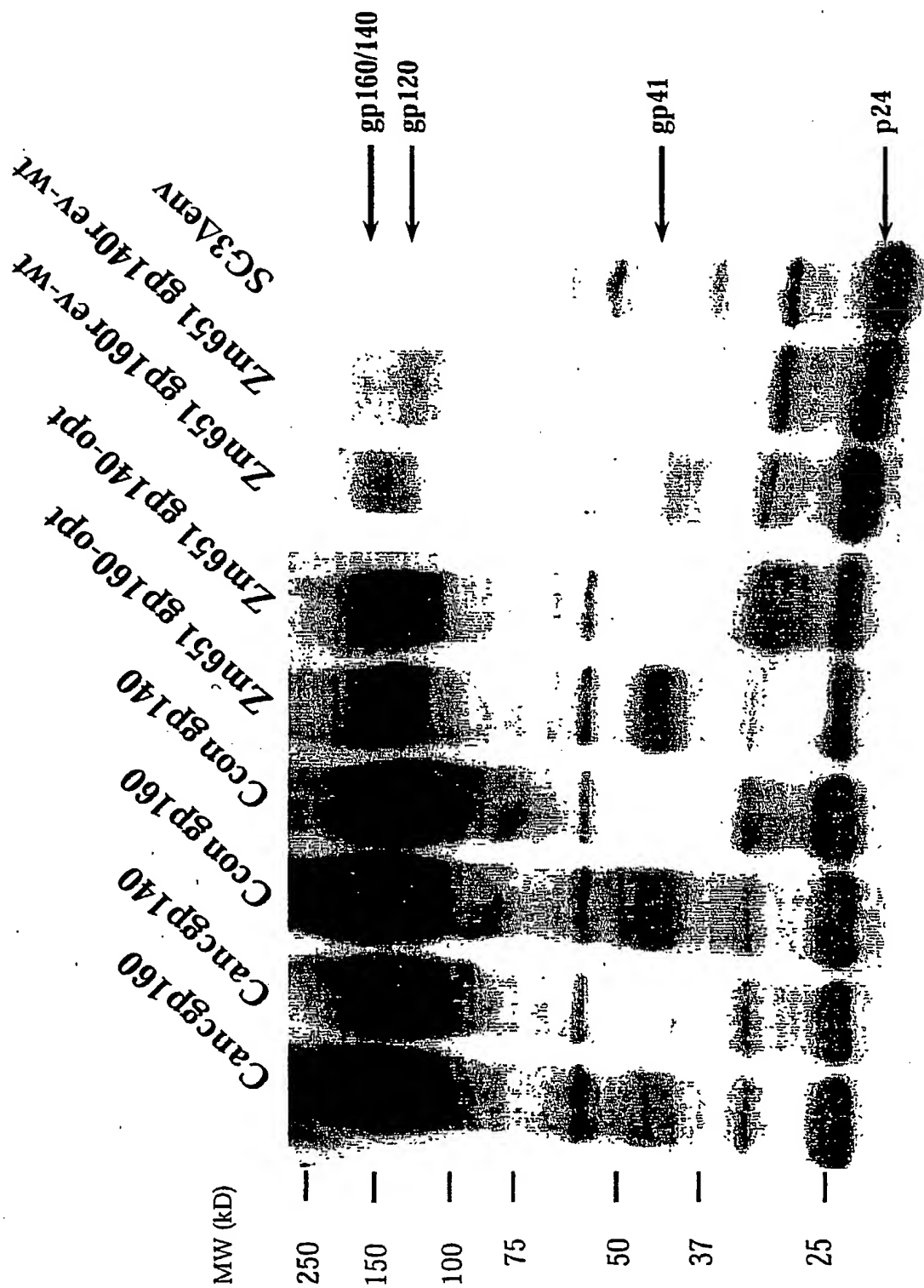


Fig. 9

15/178

Fig. 10A



16/178

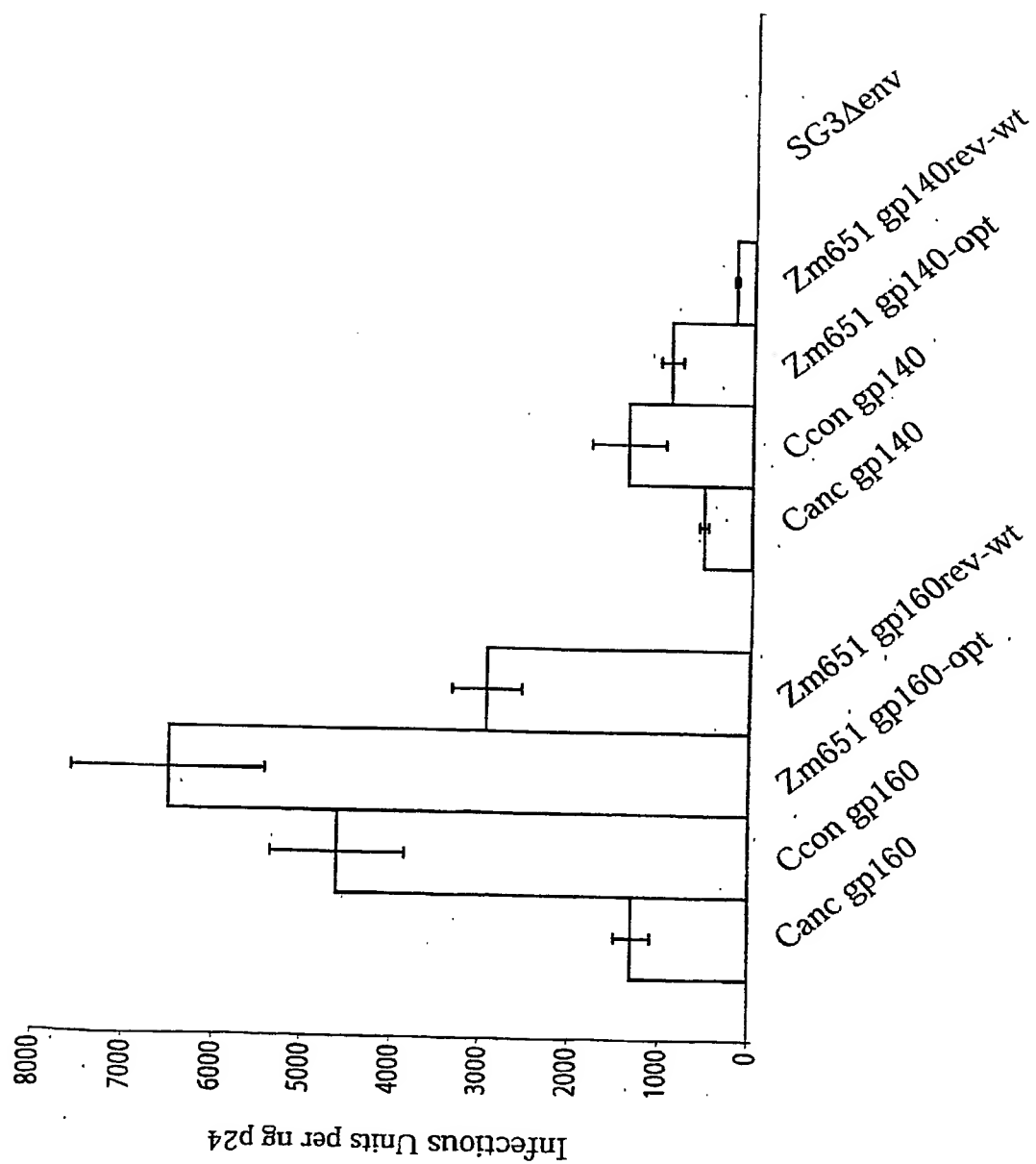
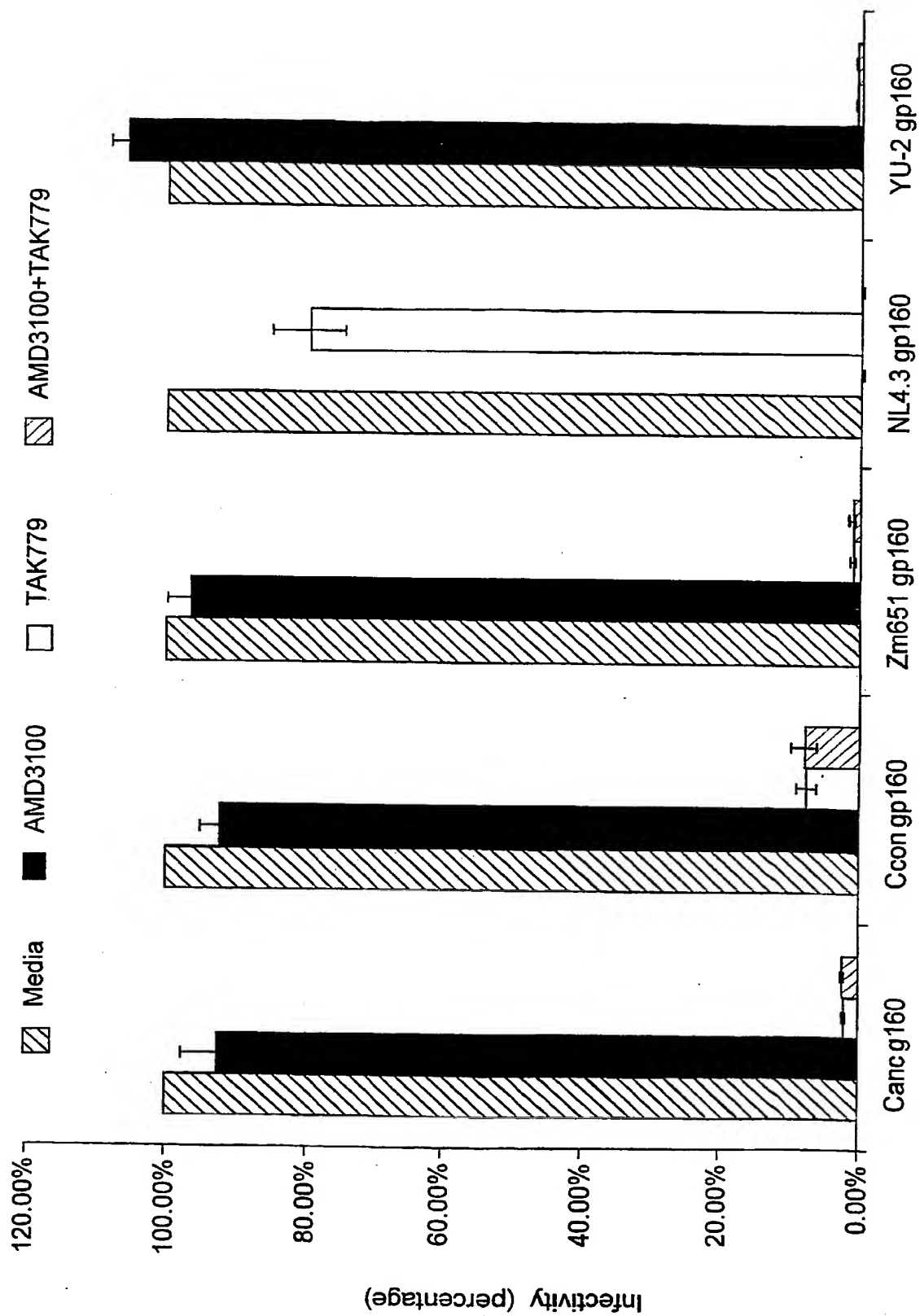


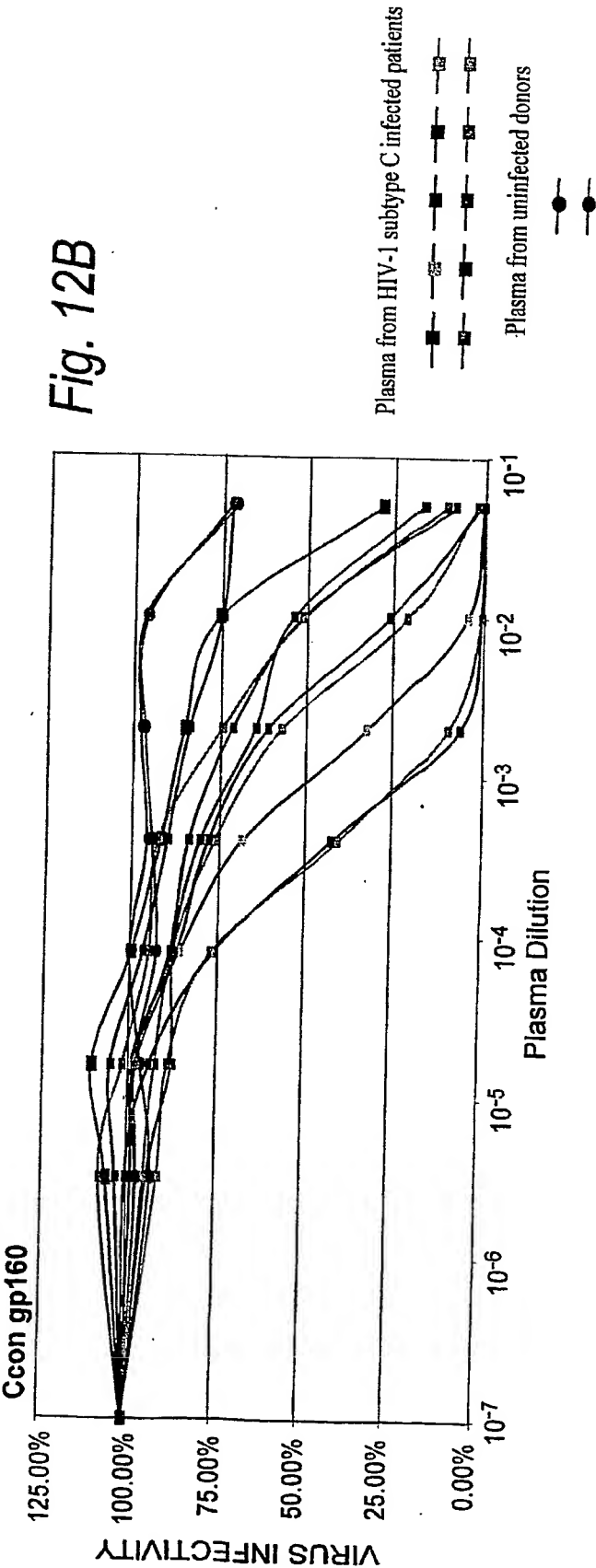
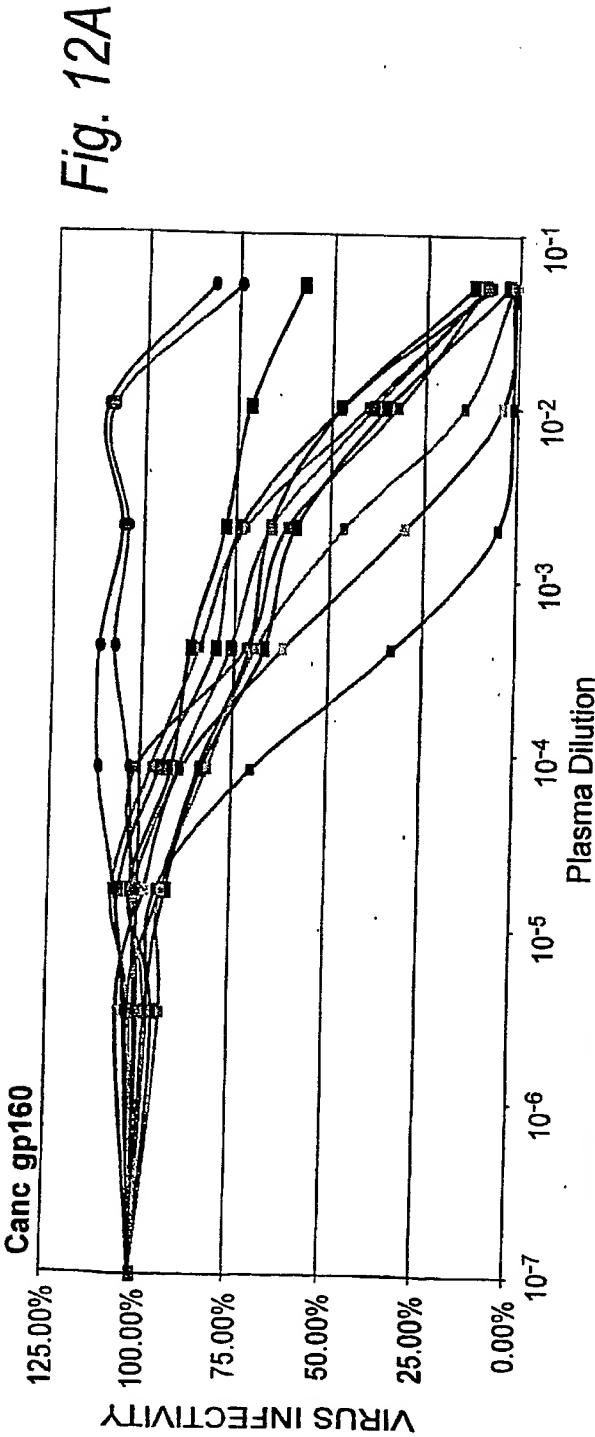
Fig. 10B



17/178

Fig. 11





19/178

Fig. 12C

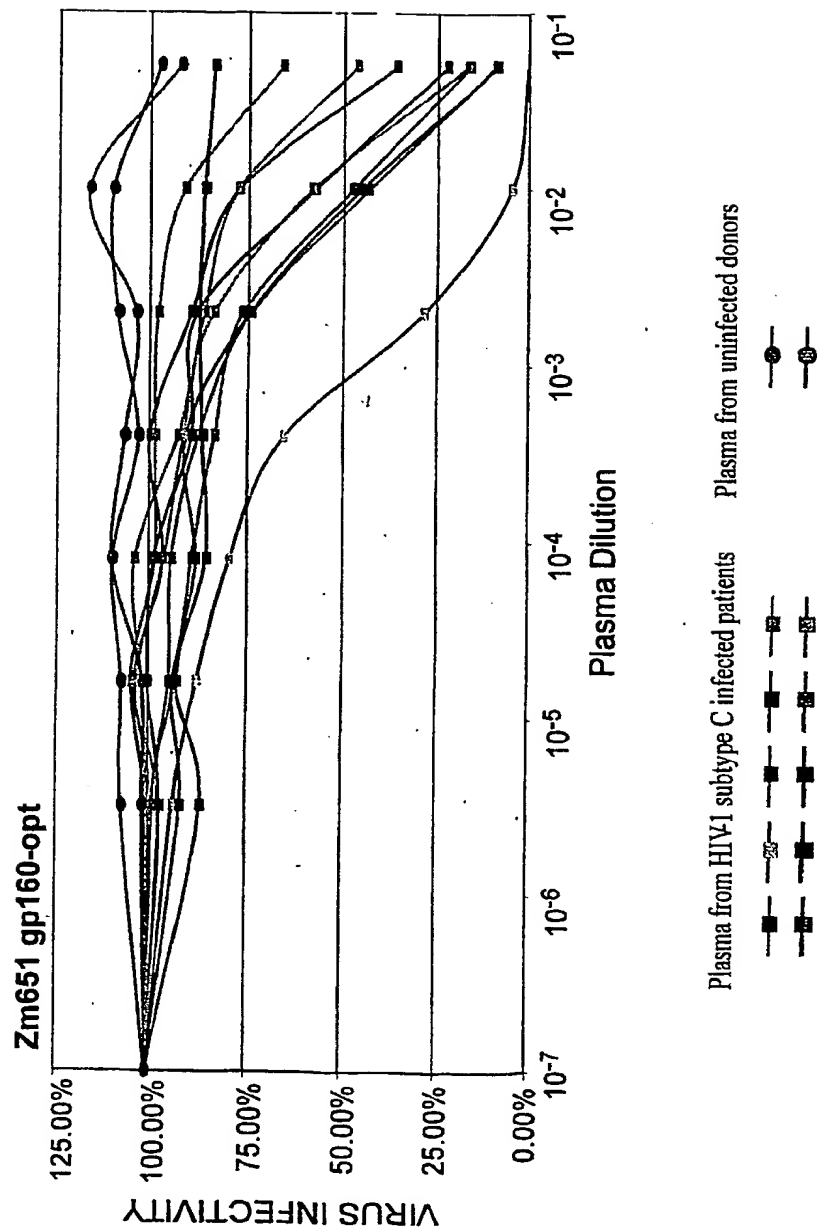


Fig. 13A

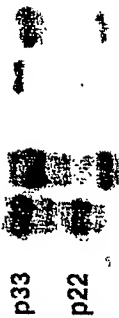
Gag

96ZM651  
Consensus C  
Mock



Nef

96ZM651  
Consensus C  
Mock



C.con.gag (subtype C con sensus gag)

MGARASILRGGKLDTWKIRLRPGGKKRYMIKHLVWASRELERFALNPGLLETSEGCKQIMKQLQPA  
LQTGTEELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEQNKSSQKTQQAEEAAGKVSQNYPI  
VQNLQGGQMVHQAI SPRTLNAWVKVIEEKAFSPVIMFTALSEGATPQDLNLTMLNTVGGHQAAMQMLKDT  
INEEAAEWDRLLHPVHAGPIAPGQMRPRGSDIAGTTSTL QEQIAWMTSNPPVPVGDYKRWIILGLNKIV  
RMYSPPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDILLVQANPDCKTILRALGPGASLE  
EMNTACQGVGGPSHKARVLAEAMSQANNTNIMMQRSNFKGPKRI VKCFNCGKEGHIARNCRAPRKKGCWK  
CGKEGHQMKDCTERQANFLGKIWP SHKGRPGNFIQSRPEPTAPAESFRFEETTPA  
PKQEPKDRPLETSLKSLFGSDPLSQ

Fig. 13C

C.con.nef (subtype C consensus nef)

MGGKWSKSSIVGWPVVRERIRRTPEAAEGVGAASQDLDKYGALTSSNTATNNADCAWLEAQEEEEEV  
GFPVRPQVPLRPMTYKAAFDLSFFLKEKGGLEGLIYSKKRQEIIDLWVYHTQGFDPDQWNYTPGPGVRY  
LTFGWCFCFLVPVDPREVEEANEENCLLHPMSQHGMEDEDEVLKWKFDSHLARHMARELHPEYYKDC

Fig. 13D

21/178

## C.con.gag (subtype C consensus gag. Not in the public domain)

GCCGCCCATGGGGCCCGCGCCAGCATCTCTGGCGGGGCAAGCTGGACACCTGGGAGAAGATCCGCC  
 TGGCCCCCGGCAAGAGCGCTACATGATCAAGCACCTGGTGTGGCCAGCCGCGAGCTGGAGCGCTT  
 CGCCCTGAACCCCGGCTGTGGAGACAGCGAGGCTGCAAGCAGATCATGAAGCAGCTGACGCCGCC  
 CTGCAGACCGGACCGAGGAGCTGCGAGCTGTACAACACCGTGGCCACCTGTACTGCTGTCACGAGA  
 AGATCGAGGTGCGGACACCAAGGAGGCTTGGACAAGATCGAGGAGGAGCAGAACAGAGCCAGCAGAA  
 GACCCAGAGGCGAGGCGCGCCGACGCGCAAGGTGAGCCAGAACTACCCATCGTGCAGAACCTGCAG  
 GGCCAGATGTTGACCCAGGCCATCAGCCCCCGCACCTGAACGCTGGGTGAAGTGATCGAGGAGAAGG  
 CCTTCAGCCCCGAGGTGATCCCATGTTTCAACGCTGAGCGAGGCGCCACCCCCAGGACCTGAAACAC  
 CATGCTGAACACCGTGGCGGCCACAGGCGGCCATGAGATGCTGAAGGACACCATCAACGAGGAGGCC  
 GCCGAGTGGGACCGCTGCACCCCGTGCACGCGCGGCCCATCGCCCCCGGCCAGATGCGGAGCCCCGCG  
 GCAGCGACATCGCCGGCACACCCAGCACCTGCAGGAGCAGATCGCTGGATGACAGCAACCCCCCCCGT  
 GCCGTGGCGACATCTAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGGCATGTACAGCCCC  
 GTGAGCATCTTGACATCAAGCAGGCGGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGA  
 CCTGCGCGCGAGCGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAAGACGC  
 CAACCCGACTGAAGAEATCTCTGGCGCCCTGGGCCCGCGCCAGCTGGAGGAGATGATGACCGCC  
 TGCCAGGCGTGGCGGCCCGAGCCCAAGGCCCGCTGTGGCCGAGGCCATGAGCCAGGCCCAACAACA  
 CCAACATCATGATGACGCGCAGCAACTTCAAGGGCCCCAAGCGCATCGTGAAGTCTTCAACTGCGGCAA  
 GGAGGGCCACATCGCCCGCAACTGCGCGGCCCGCCCGCAAGAAGGCTGCTGGAAGTGCGGCAAGGAGGC  
 CACCATGAAGGACTGCAACGAGCGCCAGGCCAACTTCTTGGGCAAGATCTGGCCAGCCACAAGGGCC  
 GCGCGGCAACTTCTGAGAGCGCCCGAGCCACCGCCCCCGCGAGAGCTTCCGCTTCGAGGA  
 GACCAACCCCGCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCCCTGACCAGCCCTGAAGAGCCCTGTTCCGGC  
 AGCGACCCCTGAGCCAGTAA

Fig. 13E

## C.con.nef (subtype C consensus nef. Not in the public domain)

GCCGCCCATGGGGCCCGCGCCAGTGGAGCAAGAGCAGCATCGTGGGCTGGCCCCCGCGTGGCGAGCGCATCC  
 GCCGACCGAGCCCGCGCGAGGGCGTGGGCGCGCCAGCAGGACCTTGAACAAGTACGGCGCCCTGAC  
 CAGCAGCAACCGCCACCAACAACCGCGACTGCGCTGGCTGGAGGCCAGGAGGAGGAGGAGGTG  
 GGCTTCCCCGTGGCCCCCAGGTGCCCTTGGCGCCCATGACCTACAAGGCCGCTTCCGACCTGAGCTTCT  
 TCCTGAAGGAGAAAGGCGGCCCTGGAGGGCTGATCTACAGCAAGAAGCGCCAGGAGATCCTGGACCTGTG  
 GGTGTACCACACCCAGGGCTTCTTCCCGCATGGCAGAACTACACCCCGGCCCGCGTGGCTGCGCTACCCC  
 CTGACCTTCGGCTGGTGTCTCAAGCTGGTCCCGTGGAACCCCGGAGGTGGAGGAGGCCAACGAGGGCG  
 AGAACAACTGCTGTGACCCCATGAGCCAGCACCGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTG  
 GAAGTTCGACAGCCACCTGGCCCCCGGCCACATGGCCCCGCGAGCTGCACCCCCGAGTACTACAAGGACTGC  
 TGA

Fig. 13F

22/178

CONs.env (group M consensus env gene. This one contain the consensus sequence for variable regions in env gene)

MRVRGIQRNCQHLMRWGTLILGMLMI CSAENLWTVYGVVWKEANTLFCASDAKAYDTEVHNV  
WATHACVPTDPNPQEIIVLENTFNFMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVNVTN  
TTNNTTEKGEIKNCSFNIITEIRDKKQVYALFYRLDVVPIDDNNNNSSNYRLINCNTSAITQACPVSF  
EPIPIHYCAPAGFALLKCNCKKENGTPCKNVSTVQCTHGKIPVSTQLLNGSLAEEIIIRSENI TNN  
AKTIIIVQLNESVEINCTRPNNNTNRKSIRIGPGQAFYATGDIIGDIRQAHCNISGKWNKTLOQVAKKLRE  
HFNNKTIIFKPSSGGDEIITHSFNCRGEFFYCNTSGLFNSWIGTKNNNTNDTI TLPCR IKQI INM  
WQGVQAMYPPIEGKITCKSNI TGLLLTRDGGNNNTNETEIFRPGGDMRDNRSELYKYKVVKIEPLG  
VAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVOARQLLSGIVQQSNLLRAIEAQHLL  
LQLTWVGIKQLQARVLAVERYLKDQQLLGIWCSGKLCITTTVPWNSSWSNKSQDEIWDNNMTWMEWEREI  
NNYTDIIYSLIEESQOQKNEQELLALDKWASLWNWFDITNLWYIKIFIMIVGGLIGLRIVFAVLSIV  
NRVRQGYSPLSFQTLIPNPRGPDRPEGIEEGEGEQDRDRSIRLVNGLALAWDDLRLSLCFSYHRLRDFI  
LIAARTVELLGRKGLRRGWEALKYLWNLLQYWGQELKNSAISLLDTTAAIAVAEGTDRVIEVVQACRAIL  
NIPRRIRQGLERALL

Fig. 14A

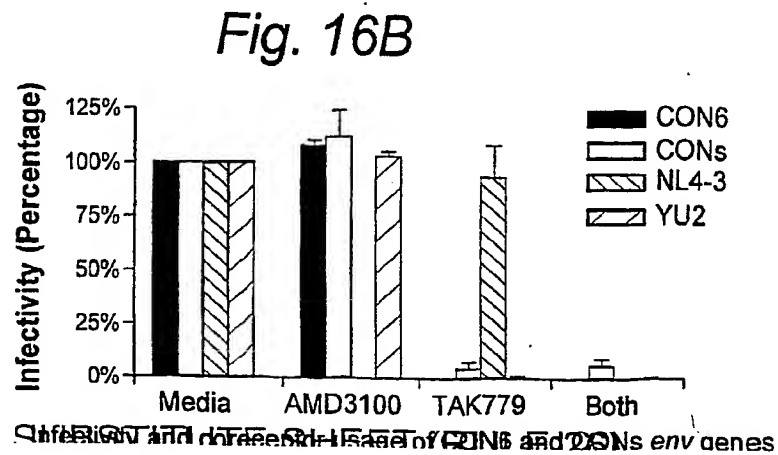
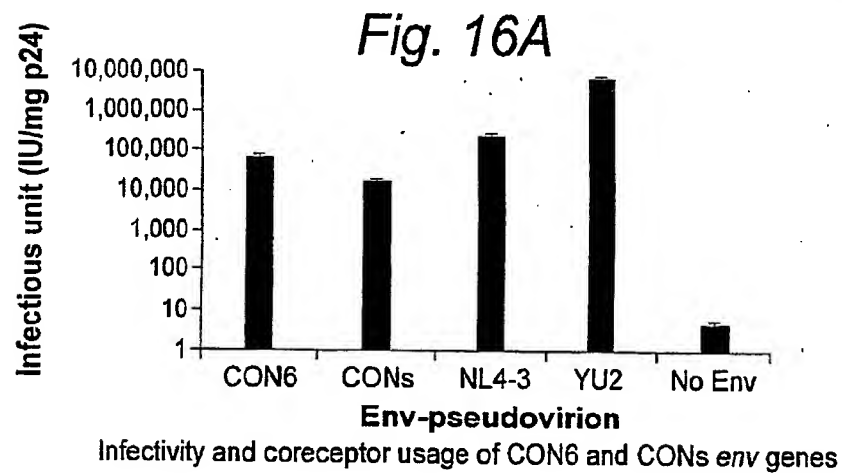
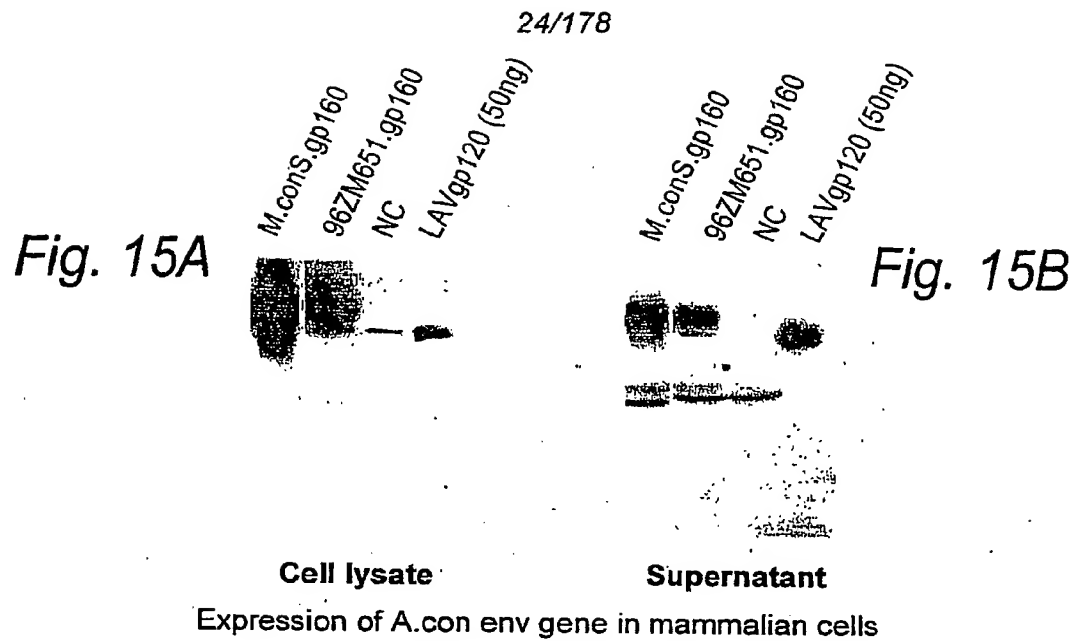


Fig. 14C

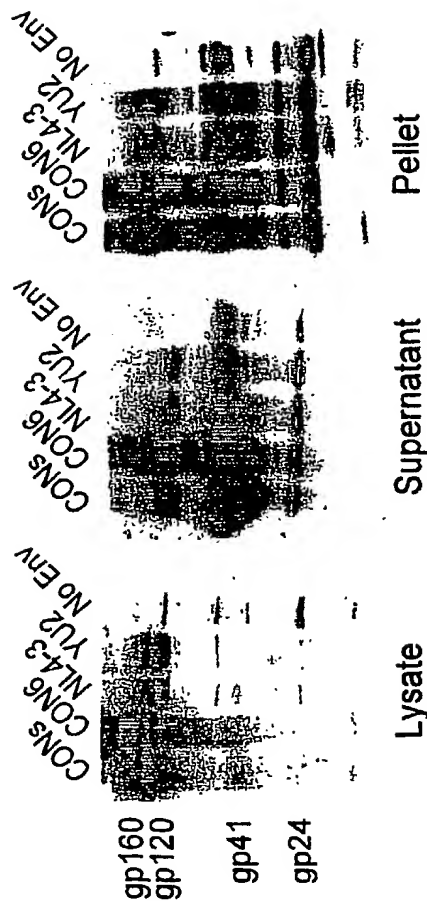
**Fig. 14B**

CONs.env (gorup M consensus env gene. This one contain the consensus sequence for variable regions in env gene. The identical amino acid sequences as in the public domain)

GCCGCCGCCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTG  
GCGCTGGGGCACCTGATCCTGGGCATGCTGATGATCTGCTCCGCCGCCG  
AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC  
AACACCACCTGTTCTGCGCTCCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCC  
AGGAGATCGTGCTGGAGAACGTGACCGAGAAGTTCAACATGTGGAAGAAC  
AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTC  
CCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAACTGCA  
CCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAG  
ATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCA  
GAAGGTGTACGCCCTGTTCTACCGCTGGACGTGGTGCCCATCGACGACA  
ACAACAACAACTCCTCCAACCTACCGCTGATCAACTGCAACACCTCCGCC  
ATCACCCAGGCC TGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTA  
CTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCA  
ACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAAGTGCAACCCACGGC  
ATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA  
GGAGGAGATCATCATCCGCTCCGAGAACATCACCACAACGCCAAGACCA  
TCATCGTGACGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAAC  
AACAACACCCGCAAGTCCATCCGCATCGGCCCGGCCAGGCCCTTCTACGC  
CACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCG  
GCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAGAAGCTGCGCGAG  
CACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGGCGACCT  
GGAGATCACCAACCACTCCTTCAACTGCCCGGGCGAGTTCTTCTACTGCA  
ACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAAC  
AACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT  
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCG  
AGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGC  
GACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCGGCCG  
CGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTGG  
TGAAGATCGAGCCCTGGGCGTGCGCCCCACCAAGGCCAAGCGCCGCGTG  
GTGGAGCGCGAGAAGCGCGCGCTGGGCATCGGCGCCGTGTTCTGGGCTT  
CCTGGGCGCCGCGGCTCCACCATGGGCGCCGCTCCATCACCTGACCG  
TGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGACGAGCAGTCCAACCTG  
CTGCGCGCCATCGAGGCCAGCAGCACCTGCTGACGCTGACCGTGTGGGG  
CATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACCTGAAGG  
ACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCACC  
ACCACCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAGGACGAGAT  
CTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAATAACA  
CCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACAGCAGGAGAAG  
AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACCTG  
GTTTCGACATCAACAACTGGCTGTGGTACATCAAGATCTTCATCATGATCG  
TGGCGCGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCCATCGTG  
AACC GCGTGCGCCAGGGCTACTCCCCCTGTCTTCCAGACCCCTGATCCC  
CAACCCCGCGGCCCGGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCG  
AGCAGGACCGCGACCGCTCCATCCGCCTGGTGAACGGCTTCTGGCCCTG  
GCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCG  
CGACTTCATCCTGATCGCCGCCCCGACCGTGGAGCTGCTGGGCGCGAAGG  
GCCTGCGCCGCGGCTGGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAG  
TACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAC  
CGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGC  
GCGCCTGCGCGCCATCTGAACATCCCCCGCGCATCCGCCAGGGCCTG  
GAGCGCGCTGCTGCTTA







Env protein incorporation in CON6 and CONs Env-pseudovirions

Fig. 17A Fig. 17B Fig. 17C

A.con.env (subtype A consensus env)

MRVMGIQRNCQHLWRWGTMLGMIICSAENLWTVVYGVVPWKDAETTLFCASDAKAYDTEVHNV  
WATHACVPTDPNPQEIINLENVTEEFNMWKNMVEQMHTDIIISLDQSLKPCVKLTPLCVTLNCSNVNVT  
NITNITDNMKGEEKNCSEFNMVTELRDCKKQVSYLFYKLDVVQINKSNSSSQYRLINCNTSAITQACPVS  
FEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAESEVMIRSENITN  
NAKNIIVQLTKPKVINCITRPNNTKRSIRIGPGQAFYATGDIIGDIRQAHCNVSRTEWNETLQKVAKQLR  
KYFNNTKIIIFTNSSGGDLIETTHSFNCGGEFFCYNTSGLFNSTWNGTKKKNSTESNDTITLPCRIKQI  
INMWQRVGQAMYAPPIQGVIRCESNITGLLLTRDGGDNNSKNETFRPGGDMRDNRSELYKVKVVKIEP  
LGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAAASITLVQARQLLSGIVQQSNLLRAIEAQQ  
HLLKLTWGIKQLQARVLAVERYLKDQQLGIWCGSKLICITNVPMNWSNKSQSEIWDNMTWLQWDK  
EISNYTDIIYNLIEESQKQKNEQDLALDKWANLWNFEDISNWLWYIKIFIMIVGGLIGLRIVFAVLS  
VINRVQGYSPLSFQHTPNPGGLDRPGRIIEEGEGEQGRDRSIRLVSGFLALAWDDLRLSLCLFSYHRLRD  
FILIAARTVELLGHSSSLKGLRLGWEGLYLWNLLLYWGRELKSAINLLDTIAIAGWTDTRVIEIGQRI  
CRAILNIPRRIRQGLERALL

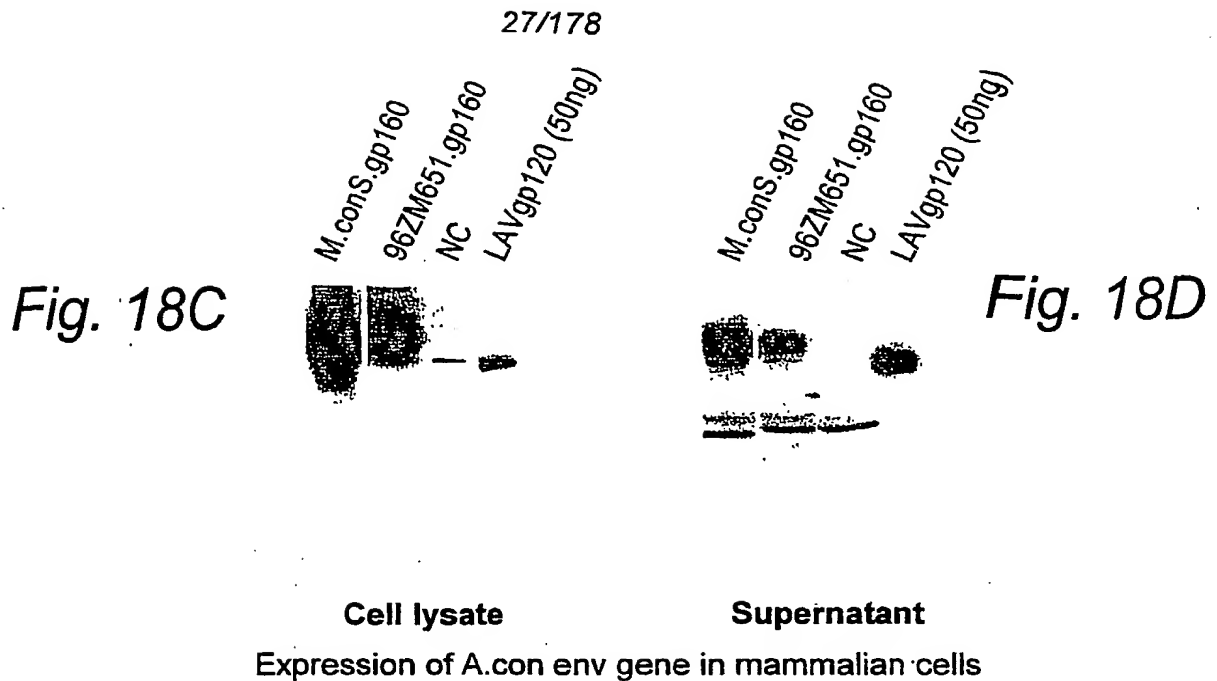
Fig. 18A

26/178

**Fig. 18B**

A.con.env (subtype A consensus env. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTG  
GCGCTGGGGCACCATGATCCTGGGCATGATCATCATCTGCTCCGCCGCCG  
AGAACCTGTGGGTGACCGTGACTACGGCGTGCCCGTGTGGAAGGACGCC  
GAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT  
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCC  
AGGAGATCAACCTGGAGAAGCTGACCGAGGAGTTCAACATGTGGAAGAAC  
AACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGGACCAGTC  
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCTGAACCTGCT  
CCAACGTGAACGTGACCACCAACATCACCACATCACCAGACAACATGAAG  
GGCGAGATCAAGAACTGCTCCTTCAACATGACCACCGAGCTGCGCGACAA  
GAAGCAGAAGGTGTACTCCCTGTTCTACAAGCTGGACGTGGTGCAGATCA  
ACAAGTCCAACCTCCTCCTCCAGTACCGCCTGATCAACTGCAACACCTCC  
GCCATCACCCAGGCCGCCCCAAGGTGTCTTTCGAGCCCATCCCCATCCA  
CTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGGAGT  
TCAACGGCACCGGCCCTGCAAGAAGCTGTCCACCGTGCAAGTGCAACCCAC  
GGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGC  
CGAGGAGGAGGTGATGATCCGCTCCGAGAACATCACCACCAACGCCAAGA  
ACATCATCGTGACGTGACCAAGCCCGTGAAGATCAACTGCACCCGCCCC  
AACACAACACCCGCAAGTCCATCCGCATCGGCCCCCGGCCAGGCCTTCTA  
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGT  
CCCCGACCGAGTGGAACGAGACCTGCGAGAAGGTGGCCAAGCAGCTGCGC  
AAGTACTTCAACAACAAGACCATCATCTTCAACCAACTCCTCCGCGCGCA  
CCTGGAGATCACCAACCACTCCTTCAACTGCGGCGGCGAGTTCTTCTACT  
GCAACACCTCCGGCCTGTTCAACTCCACCTGGAAACGGCAACGGCACCAAG  
AAGAAGAACTCCACCGAGTCCAACGACACCATCACCTGCCCTGCCGCAT  
CAAGCAGATCATCAACATGTGGCAGCGCGTGGGCCAGGCCATGTACGCCC  
CCCCCATCCAGGGCGTGATCCGCTGCGAGTCCAACATCACCAGCCTGCTG  
CTGACCCGCGACGGCGGC GACAACAACCTCCAAGAACGAGACCTTCCGCCC  
CGGCGGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTACA  
AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAGGCCAAGCGC  
CGCGTGGTGGAGCGCGAGAAGCGCGCGTGGGCATCGGCGCGCTGTTCCT  
GGGCTTCTGGGCGCGCCGCTCCACCATGGGCGCGCCCTCCATCAACC  
TGACCGTGACGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCC  
AACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGT  
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGGCCGTGGAGCGCTACC  
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC  
TGCACCACCAACGTGCCCTGGAACCTCCTCCTGGTCCAACAAGTCCCAGTC  
CGAGATCTGGGACAACATGA CCTGGCTGCAGTGGGACAAGGAGATCTCCA  
ACTACACCGACATCATCTACAACCTGATCGAGGAGTCCCAGAACCCAGCAG  
GAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGGCAACCTGTG  
GAACTGGTTCGACATCTCCAACCTGGCTGTGGTACATCAAGATCTTCATCA  
TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCC  
GTGATCAACCGCGTGGCGCCAG GGCTACTCCCCCTGTCTTCCAGACCCA  
CACCCCCAACCCCGGCGGCCTGGACCGCCCCGGCCGCATCGAGGAGGAGG  
GCGGCGAGCAGGGCCGCGACCGCTCCATCCGCTGGTGTCCGGCTTCCCTG  
GCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCG  
CTGCGCGACTTCATCCTGATCGCCGCCCCGCACCGTGGAGCTGCTGGGCC  
ACTCCTCCCTGAAGGGCCTGCG CCTGGGCTGGGAGGGCCTGAAGTACCTG  
TGGAACCTGCTGCTGTACTGGGGCCGCGAGCTGAAGATCTCCGCCATCAA  
CCTGCTGGACACCATCGCCATCGCCGTGGCCGGCTGGACCGACCGCGTGA  
TCGAGATCGGCCAGCGCATCTGCCGCGCCATCCTGAACATCCCCCGCGC  
ATCCGCCGAGCTGAGTCCGATCTGCTGAA

*Fig. 19A*

M.con.gag (group M consensus gag. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGA  
CGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCC  
TGAAGCACCTGGTGTGGGCTCCCGCGAGCTGGAGCGCTTCGCCCTGAAC  
CCCGGCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCGGCCAGCT  
GCAGCCCCGCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA  
CCGTGGCCACCCTGTACTGCGTGCAACAGCGCATCGAGGTGAAGGACACC  
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAA  
GACCCAGCAGGCCGCCGCCGACAAGGGCAACTCCTCCAAGGTGTCCCAGA  
ACTACCCCATCGTGCAAGCCTGCAGGGCCAGATGGTGCACCAGGCCATC  
TCCCCCGCACCCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTT  
CTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC  
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC  
ATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCG  
CCTGCACCCCGTGCACGCCGGCCCCATCCCCCGGCCAGATGCGCGAGC  
CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC  
GCCTGGATGACCTCCAACCCCCCATCCCCGTGGGCGAGATCTACAAGCG  
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCGTGT  
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCGCGCACTACGTG  
GACCGCTTCTTCAAGACCTGCGCGCCGAGCAGGCCACCCAGGACGTGAA  
GAACTGGATGACCGACACCTGCTGGTGCAGAACGCCAACCCCGACTGCA  
AGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCCTGGAGGAGATGATG  
ACCGCCTGCCAGGGCGTGGGCGGCCCGGCCACAAGGCCCGCGTGTGGC  
CGAGGCCATGTCCAGGTGACCAACGCCGCCATCATGATGCAGCGCGGCA  
ACTTCAAGGGCCAGCGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAG  
GGCCACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAA  
GTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA  
ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCGGCAACTTC  
CTGCAGTCCCCGCCCGAGCCACCGCCCCCCCCCGCCGAGTCCTTCGGCTT  
CGGCGAGGAGATACCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGC  
CCCCCTGACCTCCCTGAAGTCCCTGTTCCGCAACGACCCCTGTCCAG  
TGA

M.con.pol.nuc

28/178

Fig. 19B

GCCGCCGCATGCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACCAT  
 CAAGATCGGCGGCCAGCTGAAGGAGGCCCTGCTGGCCACCGGCGCCGACG  
 ACACCGTGCTGGAGGAGATCAACCTGCCCGCAAGTGGAAGCCCAAGATG  
 ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT  
 GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGGTGGGCCCCA  
 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACC  
 CTGAACTTCCCCATCTCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCC  
 CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGA  
 TCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC  
 TCCAAGATCGGCCCCGAGAACCCTTACAACACCCCCATCTTCGCCATCAA  
 GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  
 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCC  
 GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC  
 CTACTTCTCCGTGCCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCA  
 CCATCCCCTCCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAAC  
 GTGCTGCCCCAGGGCTGGAAGGGCTCCCCGCCATCTTCCAGTCCCTCCAT  
 GACCAAGATCCTGGAGCCCTTCCGCACCCAGAACCCCCGAGATCGTGATCT  
 ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG  
 CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGCGCTGGGGCTT  
 CACCACCCCCGACAAGAAGCACCAAGAAGGAGCCCCCTTCTGTGGATGG  
 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCC  
 GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT  
 GAACTGGGCCTCCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCA  
 AGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCTGACCGAG  
 GAGGCCGAGCTGGAGCTGGCCGAGAACC CGGAGATCCTGAAGGAGCCCGT  
 GCACGGCGTGTTACTACGACCCCTCCAAGGACCTGATCGCGAGATCCAGA  
 AGCAGGGCCAGGACCAAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG  
 AACCTCAAGACCGGCAAGTACGCCAAGATGCGCTCCGCCACACCAACGA  
 CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGTCCATCG  
 TGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACC  
 TGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATTCCCGAGTG  
 GGAGTTCTGTGAACACCCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA  
 AGGAGCCCATCGCCGGCGCGAGACCTTCTACGTGGACGGCGCCGCCAAC  
 CGCGAGACCAAGCTGGGCAAGGCCGGCTACGTGACCGACCGCGGCCGCCA  
 GAAGGTGGTGTCCCTGACCGAGACCACCAACCAGAAAACCGAGCTGCAGG  
 CCATCCACCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC  
 GACTCCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGA  
 GTCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG  
 TGTACCTGTCTGGGTGCCCGCCACAAGGGCATCGGCGGCAACGAGCAG  
 GTGGACAAGCTGGTGTCCACCGGCATCCGCAAGGTGCTGTTCTTGACGG  
 CATCGACAAGGCCAGGAGGAGCACGAGAAGTACCACTCCAAGTGGCGCG  
 CCATGGCCTCCGACTTCAACCTGCCCCCATCGTGGCCAAGGAGATCGTG  
 GCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGT  
 GGACTGCTCCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA  
 AGATCATCCTGGTGGCCGTGCACGTGGCCTCCGGCTACATCGAGGCCGAG  
 GTGATCCCCGCGGAGACCGGCCAGGAGACCGCCTACTTCATCCTGAAGCT  
 GGCCGGCCGCTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAAC  
 TCACCTCCGCGCCCGTGAAGGCCCGCTGCTGGTGGGCGGCATCCAGCAG  
 GAGTTCCGCATCCCCTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCAT  
 GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGCGCGACCAAGGCCGAG  
 ACCTCAAGACCGCCGTGCAGATGGCCGTGTTTCATCCACAACCTTCAAGCGC  
 AAGGGCGGCATCGGCGGCTACTCCGCCGGCGAGCGCATCATCGACATCAT  
 CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCC  
 AGAACTTCCGCGTGTTACTACCGCGACTCCCGCGACCCCATCTGGAAGGGC  
 CCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGATCCAGGACAA  
 CTCGACATCAAGGTGGTGGCCCGCGCAAGGCCAAGATCATCCGCGACT  
 ACGGCAAGCAGATGGCCGGGACGACTGCGTGGCCGGCCCGCCAGGACGAG  
 GAGCTA

**Fig. 19C**

GCCGCGCCCATGGGCGGCAAGTGCTCCAAGTCCCTCATCGTGGGCTGGCC  
 CGCGGTGCGCGAGCGCATCCGCCGCACCCCGCCGCCGAGGGCGTGG  
 GCGCCGTGTCAGAGACCTGGACAAGCAGGCGCCATCACCTCTCCAA  
 ACCGCGCCCAACAACC CGACTGCGCCTGGCTGGAGGCCAGGAGGAGGA  
 GGAGGAGGTGGGCTTC CCGTGCGC CCCAGGTGCCCCCTGCGCCCA TGA  
 CCTACAAGGCCGCCCTGGACCTGTC CCACTTCTGAAGGAGAAGGCGGC  
 CTGGAGGGCTGATCTACTCCAAAGAAGCGCCAGGAGATCCTTGACCTGT  
 GGTGTACCAACCCAGGGCTACTTC CCCGACTGGCAGAACTACACCCCG  
 GCCCGGCGATCCGCTA CCCCCTGACCTTCGGCTGGTGCTTCAAGCTG  
 GTGCCCTGGACCCCGAGGAGGTGAGGAGGCCAACGAGGGCGAGAACA  
 ACTCCCTGTGCAACCCATGTGCGACGCAGGCATGGAGGACGAGGAGCG  
 CGAGGTGCTGATGTGGAAGTTGACTCCCGCTGGCCCTGCGCCATCGCC  
 CGCGAGCTGCACCCGAGTACTACAAGGACTGCTAA

**Fig. 19D**

GCGCCGCCCATGCCCCAGATCACCCCTGTGGCAGCGCCCCCTGGTGTCCAT  
CAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGCCACCGGCGCCGACG  
ACACCGTGCTGGAGGAGATCAACCTGCCC GGCAAGTGGAAAGCCCAAGATG  
ATCGGCGGCATCGGCGGCCTTCATCAAGGTGCGCCAGTACGACCAGATCCT  
GATCCGAGATCTGCGGCCAAGAAGGCCATCGGCCACCGTGCTGGTGGGCCCCA  
CCCCCGTGAAACATCATCGGCCGCAACATGTGACCCAGCTGGGCTGCACC  
CTGAACTTCCCCATCTCCCCATCGAGACCGTGCCCGCTGAAGCTGAAGCC  
CGGCATGGACGGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGA  
TCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATC  
ACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCCGTGTTGCCATCAA  
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCC  
GCCGGCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC  
CTACTTCTCCGTGCCCTGGACGAGGGCTTCCGCAAGTACACCGCCTTCA  
CCATCCCCTCCATCAAACAGAGACCCCCGGCATCCGCTACCAGTACAAC  
GTGCTGCCCGCAGGGCTGAAGAGGCTCCCCGCCATCTTCAGTCTCCAT  
GACCAAGATCCTGGAGCCTTCCGCGCCCCAGAACCCCGAGATCGTGATCT  
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG  
CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGAAGTGGGGCTT  
CACCACCCCCGACAAGAAGCACCAAGAAGGAGCCCCCTTCTGTGGATGG  
GCTACGAGCTGCACCCCGACAAGTGGACCGTGACGCCCATCCAGCTGCC  
GAGAAGGACTCTTGACCGGTGAACGACATCCAGAAGCTGGTGGGCAAGCT  
GAACTGGGCTCCCGAGATCTACCCCGGCATCAAGTGCGCCAGCTGTGCA  
AGCTGTGCGCGCGCCAAGCCCTGACCGACATCGTGCCTGACCGAG  
GAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGT  
GCACGGCGTGTACTACGACCCCTCAAGGACCTGATCGCCGAGATCCAGA  
AGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG  
AACCTCAAGACCGGCAAGTACGCCAAGATGCGCACCGCCACACCAACGA  
CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGTCCATCG  
TGATCTGGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAGGAGACC  
TGGGAGACCTGGTGGACCGACTCTGGCAGGCCACCTGGATTCCCGAGTG  
GGAGTTCTGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA  
AGGAGCCCTGGCTGTTGTGAGAGTCTTATGTCGATGCGGCCGCCAAC

30/178

CGCGAGACCAAGATCGGCAAGCGCGGCTACGTGACCGACGCGGCGCCCA  
 GAAGATCGTGTCCCTGACCGAGACCAACCAAGAAACCGAGCTGCAGG  
 CCATCCAGCTGGCCCTGCAGGACTCCGGTCCGAGGTGAACATCGTGACC  
 GACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCGACAAAGTCCGA  
 GTCCGAGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGCGCG  
 TGTACCTGTCTGGGTGCCGCCCAAGGGCATCGCGGCAACGAGCAG  
 GTGACAAGCTGGTGTCTCCGGCATCCGCAAGTGTGTTCTTGGACGG  
 CATGGAAGGCCAGGAGGACGAGAGTACCACTCAACTGGCGCG  
 CCATGGCTCCGAGTTCAACTGCCCCCATCGTGGCCAAAGAGATCGTG  
 GCCTCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGACGCGCCAGGT  
 GGAATGCTCCCGGCGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA  
 AGATCATCTGGTGGCCGTGCACGTGGCTTCCGGCTACATCGAGGCCGAG  
 GTGATCCCGCGGAGACCGGCCAGGAGACCGCTACTTCACTCTGAAGCT  
 GGCGGCGCTGGCCCGTGAAGGCGCGCTGCTGGTGGCGGCGATCCAGCAG  
 TCACCTCGCGCGCGTGAAGGCGCGCTGCTGGTGGCGGCGATCCAGCAG  
 GAGTTCGGCATCCCTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCAT  
 GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGGCGGACCGAGGCCGAGC  
 ACCTCAAGACCGCGCTGCAGATGGCCGTGTTCACTCCAACTTCAAGCGC  
 AAGGCGGCATCGCGGCTACTCGCGCGGAGCGCATCATCGACATCAT  
 CGCCACCGACATCCAGACCAAGGAGCTGCAGAGCAGATCATCAAGATCC  
 AGAACTTCGGCGTGTACTACCGGACTCCCGGAGACCCCATCTGGAAGGCG  
 CCGCCAAAGCTGTGTGAAGGCGAGGGCGCGCTGTGTGATCCAGGACAA  
 CTCCGACATCAAGGTGTGCCCCCGCGCAAGGCCAAGATCATCAAGGACT  
 ACGGCAAGCAGATGGCCGGCGCGACTGCTGCGCGGCGCGCGGCGGAG  
 GACTAA

Fig. 19D (continued)

## M.con.gag (group M consensus gag)

MGARASVLSGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETSEG CKQIIGQLQPA  
 LQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSQQTQQAADKGNSSKVSQNYPIVQN  
 LQGQMVHQAI SPRTLNAWVKVIEEKAFSPV I PMFSALSEGATPQDLNLTMLNTVGGHQAAQMMLKDTINE  
 EAAEWDRLLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 SPVSIIDIRQGPKEFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNNANPDCKTILKALPGATLEEMM  
 TACQGVGGPGHKARVLAEMSQVTNAAIMMQRGNFKGQRI IKCFNCGKEGHIARNCRAPRKKGCWKCGK  
 EGHQMKDCTERQANFLGKIWPSNKGPRGNFLQSRPEPTAPAESFGFGEETPSPKQEPKDPPEPLTSLK  
 SLFGNDPLSQ

Fig. 19E

31/178

Fig. 19F

M.con.pol (group M consensus pol)  
 MPQITLWQRPLVTKIGGQLKEALLaTGADDTVLEEINLPKWKPKMIGGIGGFIKVRQYDQILIEICGK  
 KAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPIPIETVPVKLPGMDGPKVKQWPLTEEKIKALTEICTE  
 MEKEGKISKIGPENPYNTPIFAIKKQDSTKWRKLVDFRELNKRTQDFWEVQLGIPHAGLKKKSVTVLD  
 VGDAYSFVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPIFQSSMTKILEPFRTONPEIM  
 YOYMDLTVGSDLEIGQHRAKIEELREHLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKD  
 SWTVNDIQKLVGKLNWASQIYPGKVKQLCKLRGAKALTDIVPLTEAELELAENREILKEPVGHVYYD  
 PSKOLIAEIQKQGGQDQWYTYQIYQEPFKNLTKGYAKMRSATHTNDVKQLTEAVQKIATESIVWVGKTPKFR  
 LPIKETWETWYQATWIPWEPFVNTPLVKLWYQLEKEPIAGAEFTFYVDGAANRETGLGKAGYVTD  
 RGRQKVVSLTETTNQKTELOAIHLALQDSGSEVNIIVTDSQYALGIAQAPDKSESELVNQIEQLIKKEK  
 VYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFDGIDKAEHEKHYHSNWRAMASDFNLPPVAKAIVASC  
 DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPFV  
 KVIHTDNGSNFTSAVKAACWAGIQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAV  
 FIHFKRKGIGGYSAGERIIDIA TDIOTKELQKIQNFRVYRDSRDPWKGPAKLLWKGEQAV  
 IQDNSDIKVVPRRKAKIRDYGKQMGAGDDCVAGRQDED

M.con.nef (group M consensus nef)

MGGKWSKSSI VGWPAVRERIRRTHPAAEGVGAVSQDLDKHGAITSSNTAANNPDCAWLEAQEEEEVEGFP  
 VRPQVPLRPMTYKAAALDSLHFLKEKGGLGLIYSKKRQELDLWVYHTQGYFPDQWQNTYTPGPIRYPLTF  
 GWCFLVLPVDPPEEVEEANEENNSLLHPMCQHGMEDEREVLWKFDSRLALRHIARELHPYKYDC

Fig. 19G

C.con.pol (subtype C consensus pol)  
 MPQITLWQRPLVSIKVGQIKEALLaTGADDTVLEEINLPKWKPKMIGGIGGFIKVRQYDQILIEICGK  
 KAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPIPIETVPVKLPGMDGPKVKQWPLTEEKIKALTAICEE  
 MEKEGKITIGPENPYNTPVFAIKKQDSTKWRKLVDFRELNKRTQDFWEVQLGIPHAGLKKKSVTVLD  
 VGDAYSFVPLDEGFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPIFQSSMTKILEPFRQANPEIM  
 YOYMDLTVGSDLEIGQHRAKIEELREHLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKD  
 SWTVNDIQKLVGKLNWASQIYPGKVKQLCKLRGAKALTDIVPLTEAELELAENREILKEPVGHVYYD  
 PSKOLIAEIQKQGGHQQWYTYQIYQEPFKNLTKGYAKMRTAHTNDVKQLTEAVQKIAMESIVWVGKTPKFR  
 LPIKETWETWYQATWIPWEPFVNTPLVKLWYQLEKEPIAGAEFTFYVDGAANRETGLGKAGYVTD  
 RGRQKIVSLTETTNQKTELOAIHLALQDSGSEVNIIVTDSQYALGIAQAPDKSESELVNQIEQLIKKER  
 VYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAEHEKHYHSNWRAMASEFNLPPVAKAIVASC  
 DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPFV  
 KVIHTDNGSNFTSAVKAACWAGIQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAV  
 FIHFKRKGIGGYSAGERIIDIA TDIOTKELQKIQNFRVYRDSRDPWKGPAKLLWKGEQAV  
 IQDNSDIKVVPRRKAKIRDYGKQMGAGDDCVAGRQDED

Fig. 19H

32/178

*Fig. 20A*

B.con.gag (subtype B consensus gag. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCGAGCTGGA  
CCGCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACAAGC  
TGAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAAC  
CCCGGCCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCCTGGGCCAGCT  
GCAGCCCTCCCTGCA GACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA  
CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACC  
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAA  
GGCCCAGCAGGCCCGCCGCGGACACCGGCAACTCCTCCCAGGTGTCCAGA  
ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATC  
TCCCCCGCACCCCTGAACGCCTGGGTGAAGGTGGTGGAGGAGAAGGCCTT  
CTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGGCGCCACCC  
CCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC  
ATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCG  
CCTGCACCCCGTGACGCGCGGCCCATCGCCCCCGGCCAGATGCGCGAGC  
CCCGCGGCTCCGACATCGCCGGCACCACTCCACCCTGCAGGAGCAGATC  
GGCTGGATGACCAACAACCCCCCATCCCCGTGGGCGAGATCTACAAGCG  
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCACCT  
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG  
GACCGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCTCCCAGGAGGTGAA  
GAACTGGATGACCGAGAC CCTGCTGGTGCAGAACGCCAACCCCGACTGCA  
AGACCATCCTGAAGGCCCTGGGCCCCGCGCCACCCTGGAGGAGATGATG  
ACCGCCTGCCAGGGCGTGGGCGGCCCGGCCACAAGGCCCGCGTGTGGC  
CGAGGCCATGTCCCAGGTGACCAACTCCGCCACCATCATGATGCAGCGCG  
GCAACTTCCGCAACCAGCGCAAGACCGTGAAGTGCTTCAACTGCGGCAAG  
GAGGGCCACATCGCCAAGA ACTGCCGCGCCCCCGCAAGAAGGGCTGCTG  
GAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG  
CCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCGGCAAC  
TTCCTGCAGTCCCGCCCCGAGCCCAACCGCCCCCCCCGAGGAGTCTTCCG  
CTTCGGCGAGGAGACCACCAACCCCTCCCAGAAGCAGGAGCCCATCGACA  
AGGAGCTGTACCCCTGGCCTCCCTGCGCTCCCTGTTTCGGCAACGACCCC  
TCCTCCCAGTAA



33/178

*Fig. 20B*

B.con.env (subtype B consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCCCTATGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTG  
 GCGCTGGGGCACCATGCTGCTGGGCATGCTGATGATCTGCTCCGCGCCG  
 AGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC  
 ACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCCTACGACACCGAGGT  
 GCACAACTGTGTGGGCCACCCACGCCGTGCGTGCCCAACGACCCCAACCCCC  
 AGGAGGTGGTGCTGGA GAACGTGACCGAGA ACTTCAACATGTGGAAGAAC  
 AACATGGTGGAGCAGATGACGAGGACATCATCTCCCTGTGGGAC CAGTC  
 CCTGAAGCCCTGCGTGAAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCA  
 CCGACCTGAAGAACAACCTGCTGAA CACCAACTCCTCCTCGGCGAGAAG  
 ATGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAA CATCAACACCTCAT  
 CCGCGACAAGGTGCGAGAAGGAGTACGCCCTGTTCTACAAGCTGGA CGTGG  
 TGCCCATCGACAACAACAACAACACCTCCTACCGCTGATCTCCTGCAAC  
 ACCTCCGTGATCACCCAGGCCGTGCCCAAGGTGTCCTTCGAGCCCATCCC  
 CATC CACTACTGCGCCCCCGCGGGCTTCGC CATCCTGAAGTGCAACGACA  
 AGAAGTTCAACGGCACCGGCCCTGCACCAACGTGTCCACCGTGCAGTGC  
 ACCCACGGCATCCGCCCCGTGGTGTCCAACAGCTGCTGCTGAACGGCTC  
 CCTGGCCGAGGAGGAGGTGGTGATCGCTCCGAGA ACTTCACCGACAACG  
 CCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACC  
 CGCCCAACAACAACAACCGCAAGTCCATCCACATCGGCCCGGGCCGCGC  
 CTTCTACACCAACGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCA  
 ACATCTCCCGCGCCAAGTGGAAACAACCC TGAAGCAGATCGTGAAGAAG  
 CTGCGCGAGCAGTTCTGCAACAAGACCA TCGTGTTCAACCAGTCCCTCCG  
 CGGCGACCCCGAGATCGTGATGCACTCCTTCAACTGCGGGCGGCGAGTTCT  
 TCTACTGCAACACCACCCAGCTGTTCAA CTCCACCTGGAA CGACAACGGC  
 ACCTGGAACAACACCAAGGACAAGAACA CCATCACCTGCTCCTGCCG CAT  
 CAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTACGCC  
 CCCC CATCCGCGGCCAGATCCGCTGCTCCTCCAACATCACCGGCTGCTG  
 CTGACCCGCGACGGCGGCAACAACAACAACGACACCGAGATCTTTCGCC  
 CGGCGGCGGCGACATGCGCGACAAC TGGCGCTCCGAGCTGTACAAGTACA  
 AGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCACCAAGGCCAAGCGC  
 CGCGTGGTGAGCGCGAGAAGCGCGCGTGGGCATCGGCGCCATGTTCT  
 GGGCTTCTGGGCGCCGCCGCTCCACCATGGGCGCCGCTCATGACCC  
 TGACCGTGAGGCCCGCCAGCTGCTGTCGGCATCGTGCAAGCAGCAGAAC  
 AACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGT  
 GTGGGGCATCAAGCAGCTGCAGGCCCGCTGCTGGCCGTGGAGCGCTACC  
 TGAAGGACCAGCAGCTGCTGGGCATCTGGGCTGCTCCGGCAAGCTGATC  
 TGCAACACCACCGTGCCCTGGAA CGCCTCTGGTCCAACAAGTCCCTGGA  
 CGAGATCTGGGAACAATGACCTGGATGGAGTGGGAGCGCGAGATCGACA  
 ACTACACCTCCCTGATCTACACCCTGATCGAGGAGTCCCAAGACCAGCAG  
 GAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCCTCCCTGTG  
 GAACTGGTTTGACATCACCAACTGGCTGTGTGATCATCAAGATCTTCA TCA  
 TGATCTGTTGGCGGCTGATCGGCTGCGCATCGTGTTGCGCTGCTGTCC  
 ATCGTGAAACCGCGTGCGCAGGGCTACTCCCCCTGTCTCTCAGACCCG  
 CCTGCCCGCCCCCGCGGCCCGACCGCCCGAGGGCATCGAGGAGG  
 GCGGCGAGCGCGACCGCGACCGCTCCGGCCGCTGGTGGACGGCTTCTG  
 GCCCTGATCTGGGACGACCTGCGCTCCCTGTGCTGTTCTCTTACCAACG  
 CCTGCGCGACCTGCTGCTGATCGTGACCGCATCGTGAGCTGCTGGGCC  
 GCCGCGGCTGGGAGGTGCTGAAGTACTGGTGAACCTGCTGCAGTACTGG  
 TCCCAGGAGCTGAAGA ACTCCGCGTGCTCCCTGCTGAACGCCACCGCCAT  
 CGCCGTGGCGAGGGCACCGACCGCGTGATCGAGGTGGTG CAGCGCGCCT  
 GCCGCGCCATCTGCA CATCCCGCGCGCATCCGCGCAGGGCTGGAGCGC  
 GCCCTGCTGTAA

33/178

*Fig. 20B*

B.con.env (subtype B consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCGCCGCCATGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTG  
 GCGCTGGGGCA CATGCTGCTGGGCATGCTGATGATCTGCTCCGC CGCCG  
 AGAAGCTGTGGGTGACCGTGTA CTACGGCGTGCCCGTGTGGAAGGAGGCC  
 ACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACC GAGGT  
 GCACAACTGTGGGCCACCCA CGCCTGCGTGCCCA CCGACCCCAA CCCCC  
 AGGAGGTGGTGTCTGGAGAACGTGACCGAGA ACTTCAACATGTGGAAGAAC  
 AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGAC CAGTC  
 CCTGAAGCCCTGCGTGAA GCTGACCC CCGTGTGCGTGACCCTGAA CTGCA  
 CCGACCTGAAGAACAA CCTGCTGAACACCAACTCCTCCTCCGCGAGAGAG  
 ATGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCA CCACCTCAT  
 CCGCGACAAGGTGCGAGAAGGAGTACGCCCTGTTCTACAAGCTGGA CGTGG  
 TGCCCATCGACAACAA CAACAACACCTCCTACCGCCTGATCTCCTGCAAC  
 ACCTCCGTGATCACCCAGGCTGCGCCCAAGGTGTCCTTCGAGCCCATCCC  
 CATCCTACTGCGCC CCGCGCGCTTCCCATCCTGAAGTGCAA CGACA  
 AGAAGTTCAACGGCACCGGCCCTGCGACCAACGTGTCCACCGTGCAGTGC  
 ACCCACGGCATCCGCC CCGTGGTGTCCA CCGAGCTGCTGCTGAACGGCTC  
 CCTGGCCGAGGAGGAGGTGGTGTCCGCTCCGAGAACTTCAACGACAACG  
 CCAAGACCATCATCGTGCAGCTGAA CGAGTCCGTGGAGATCAACTGCACC  
 CGCCCAAACAA CAACA CCGCAAGTCCATCCACATCGGCC CCGGC CGCGC  
 CTTCTACACCA CCGGC GAGATCATCGCGCATCCGCCAGGCCCA CTGCA  
 ACATCTCCCGCGCCAAGTGGAACAA CACCTGAAGCAGATCGTGAAGAAG  
 CTGCGCGAGCAGTTCGGCAACAAGACCATCGTGTTCAACCAGTCCTCGG  
 CGGCGA CCCCGAGATCTGTGATGCACTCCTTCAACTGCGGCGGCGAGTCT  
 TCTACTGCAACACCACCCAGCTGTTCAA CTCCACCTGGAA CGACAACGGC  
 ACCTGGAACAA CACCAAGGACAAGAACCAATCACCTGCGCCTGCGCGCAT  
 CAAGCAGATCATCAACATGTGCGAGAGGTGGGCAAGGCCATGTA CGCC  
 CCCC CATCCGCGGCCAGATCCGCTGCTCCTCAACATCACCGGCTGTCTG  
 CTGACCCGCGACGGCGGCAACAACAA CAGACACCGAGATGTTTCGCCC  
 CGGCGGCGGCGACATGCGCGA CAACTGGCGCTCCGAGCTGTACAAGTACA  
 AGGTGGTGAAGATCGAGCCCC TGCGTGGCCCCACCAAGGCCAAGCGC  
 CGCGTGGTG CAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCATGTTCT  
 GGGCTTCTGGGCGCCGCCGCTCCACCATGGGCGCGCCCTCATGACCC  
 TGACCGTGCAGGCCCCGCGAGCTGTGTCGGCATCGTGCAGCAGCAGAAC  
 AACCTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGCAGCTGACCGT  
 GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGAGCGCTACC  
 TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC  
 TGCAACCACCACCGTGCCTGGAA CGCCTCCTGGTCCAACAAGTCCCTGGA  
 CGAGATCTGGGA CAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACA  
 ACTACACCTCCCTGATCTACACCCTGATCGAGGAGTCCAGAACCCAGCAG  
 GAGAAGAACGAGCAGGAGCTGCTGGAGCTGACAAGTGGGCCCTCTGTG  
 GAACTGGTTT CGACATCACCAA CTGGCTGTGTACATCAAGATCTTCA TCA  
 TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCCGCTGCTGCTCC  
 ATCGTGAA CCGCGTGCGCCAGGGCTACTCC CCGCTGCTCCTTCAGACCCG  
 CCTGCCCGCCCCCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGG  
 GCGGCGAGCGCGACCGCGACCGCTCCGGCCGCTGGTGGACGGCTTCTG  
 GCCCTGATCTGGGACGACCTGCGCTCCCTGTGCTGTTCTCTTAC CACCG  
 CCTGCGCGA CTTGCTGTGATCGTGACCGCATCGTGGAGCTGCTGGGC  
 GCCGCGGCTGGGAGGTGCTGAAGTACTGGTGGAACCTGCTGCAGTACTGG  
 TCCCAGGAGCTGAAGAACTCCGCGTGTCCCTGCTGAACGCCACCGCAT  
 CGCCGTGGCGAGGGCACCGA CCGCGTGATCGAGGTGGTGCAGCGCGCCT  
 GCCGCGCCATCTGCATATCC CCGCGCATCCGCCAGGGCCTGGAGCGC  
 GCCCTGCTGTAA

34/178

Fig. 20C

**B.con.gag (subtype B consensus gag)**

MGARASVLSGGELDRWEKIRLRPGGKKKYLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQT  
 GSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSKKKAQQAADTGNSSQVSQNYPIVQNLOQ  
 QMVHQAI SPRTLNAWVVEEKAFSEVIPMFSALESGATPQDLNMTLNTVGGHQAAMQLKETINEEAA  
 EWDRLHPVHAGPIAPGQMRPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIV RMYSP  
 SILDIRQPKPEFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTAC  
 QGVGGPHKARVLAEMSQVTSATIMMQRGNFRNQRKTVKFCNCGKEGHIANKRCRAPRKKGCWKCGKEG  
 HQMKDCTERQANFLGKIWP SHKGRPCNFLOSRPEPTAPPEESFRFGEETTPSQKQEPIDKELYPLASLR  
 SLFGNDPSSQ

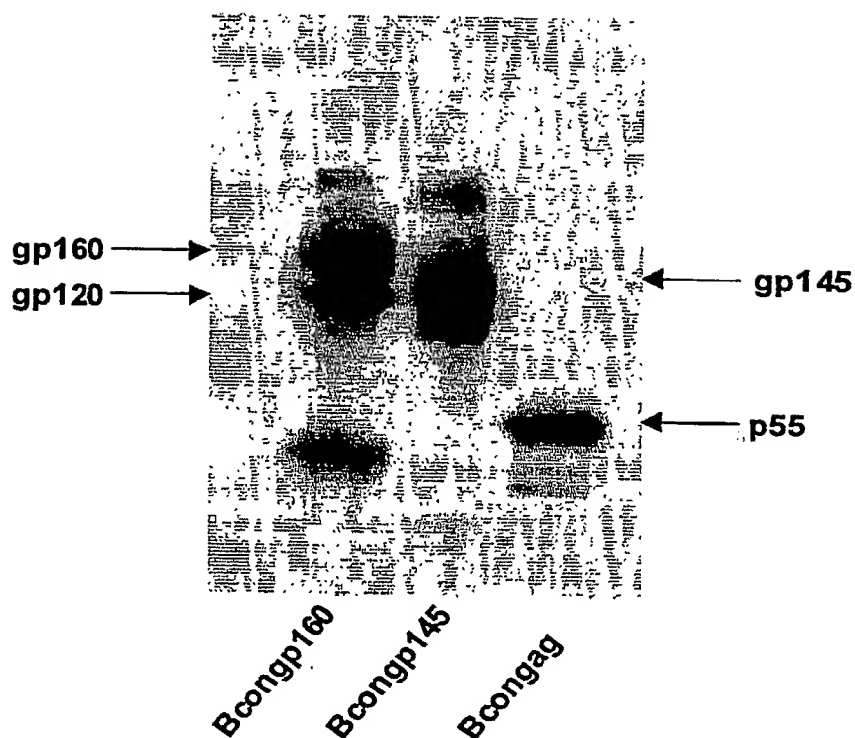
Fig. 20D

**B.con.env (subtype B consensus env)**

MRVKGIRKNYQHLLWRWGTMLLGLMLI CSAAEKLWTVV YGVVPVWKEATTLFCASDAKAYDTEVHNWAT  
 HACVPTDPNPQEVVLENTFNFMWKNMVEQMHEI ISLWDQSLKPCVKLTPLCVTLNCTDLKNNLLNT  
 NSSSGEKMEKEIKNCSEFNITTSIRDKVQKEYALFYKLDVVPIDNNNNTSYRLI SCNTSVITQACPKVSF  
 EPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLNGSLAESEVIRSENFDTN  
 AKTIIVQLNESVEINCTRPNNNTRKSIHIGPRAFYTTGEIIGDIRQAHNCISRAKWNNTLKQIVKKLRE  
 QFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTQLFNSTWNDNGTWNNTKDKNITITLPCRKQIINM  
 WQEVGKAMYAPPPIRGQIRCSSNITGLLLTRDGGNNNDTEIFRPGGDMRDNRSELYKYKVVKIEPLGV  
 APTKARRVVQREKRAVGIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGI VQQQNNLLRAIEAQHLL  
 QLTWGIKQLQARVLAVERYLKDQQLLGIWCGSKLI CTTTVPWNASWSNKSLEIWDNMTWMEWEREID  
 NYTSLIYTLIEESNQOEKNEQELLELDKWSLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVN  
 RVRQGSPLSFQTRLPA PRGPDRPEGIEEGGERDRDRSGRLVDGFLALIWDRLSLCLFSYHRLRDL  
 IVTRIVELLGRRGWEVLKYWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQACRAILHPRR  
 IRQGLERALL

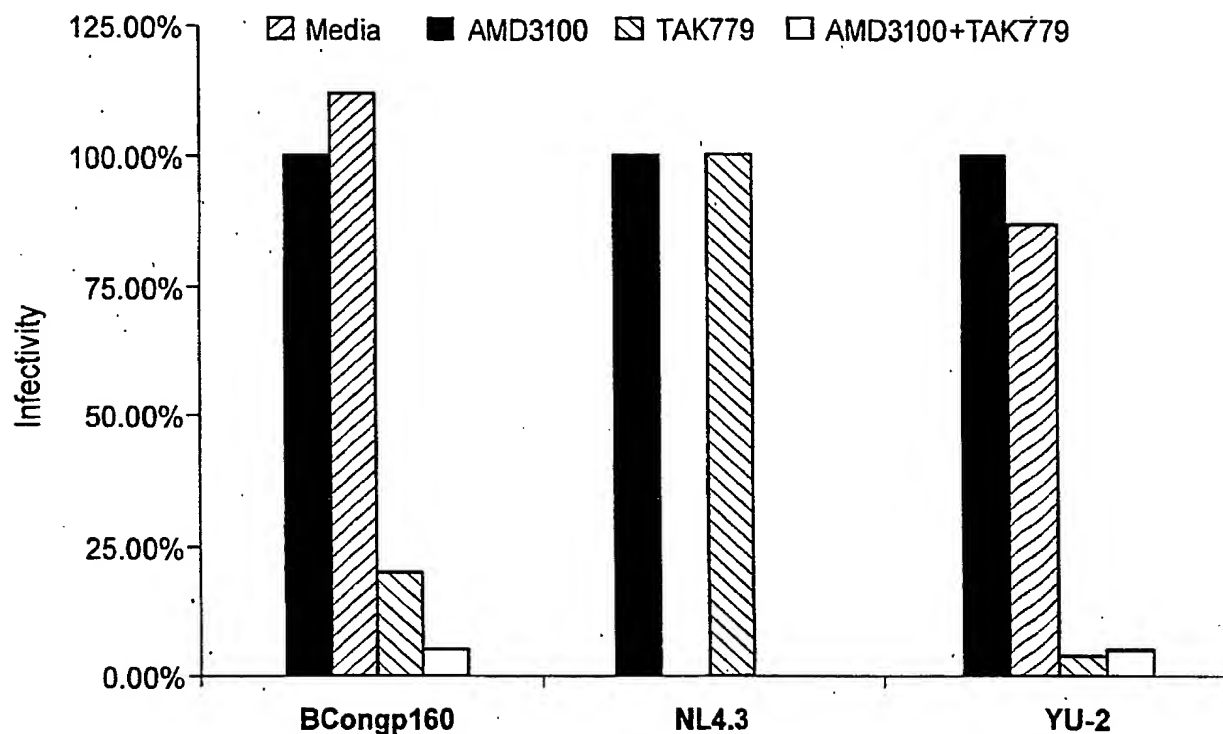
35/178

Fig. 21



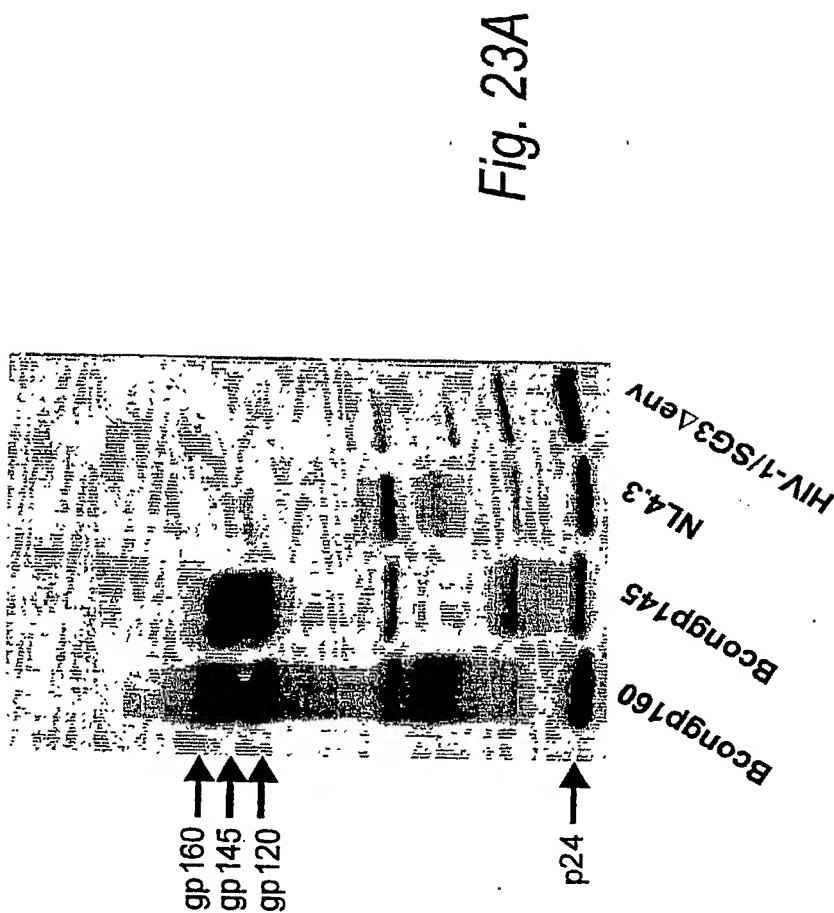
Expression of subtype B consensus *env* and *gag* genes in 293T cells. Plasmids containing codon-optimized subtype B consensus *gp160*, *gp140*, and *gag* genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2  $\mu$ g of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.

36/178

*Fig. 22***Co-receptor usage of subtype B consensus envelopes.**

Pseudotyped particles containing the subtype B consensus gp160 Env were incubated with DEAE-Dextran treated JC53-BL cells in the presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-receptor usage. NL4.3, an isolate known to utilize CXCR4 and YU-2, a known CCR5-using isolate, were included as controls.

37/178

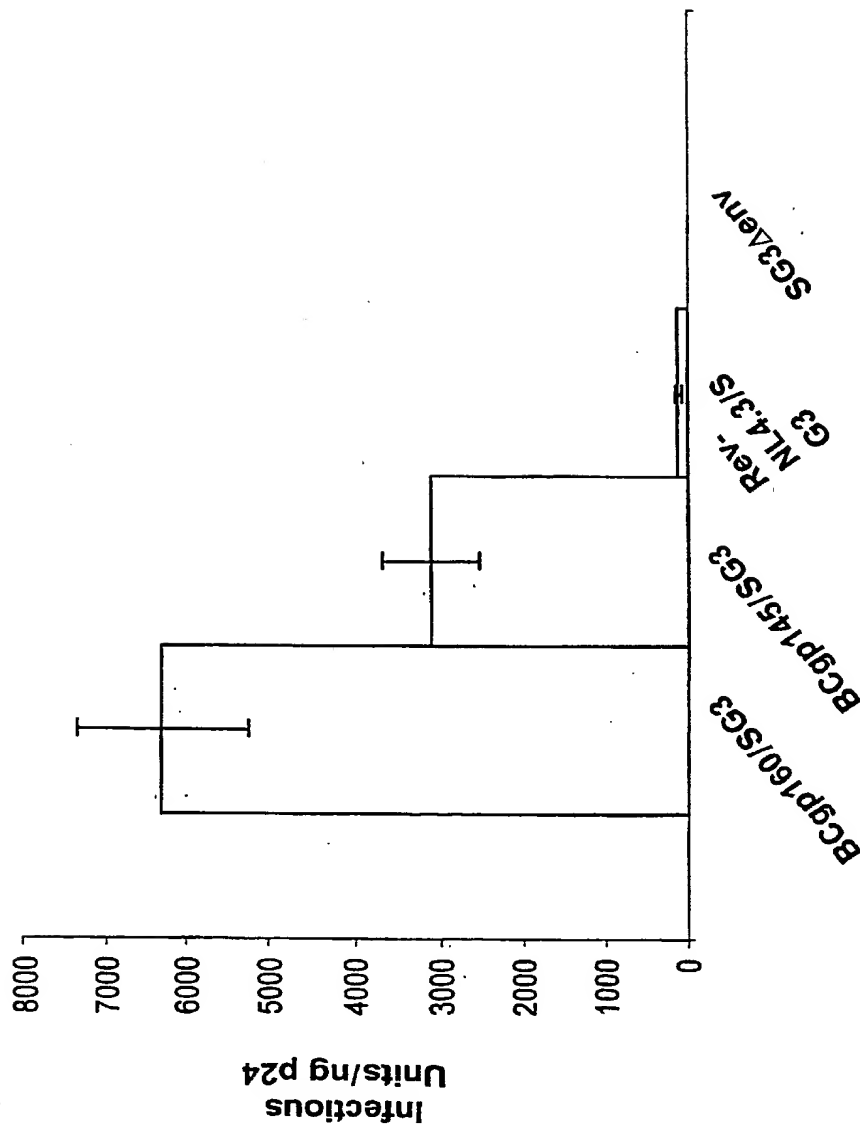


**Trans complementation of env-deficient HIV-1 with codon-optimized subtype B consensus gp160 and gp140 genes.**

Plasmids containing codon-optimized, subtype B consensus gp160 or gp140 genes were co-transfected into 293T cells with an HIV-1/SG3Δenv provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a tabletop centrifuge, filtered through a 0.2 μM filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. Trans complementation with a rev-dependent NL4.3Δenv was included for control.

38/178

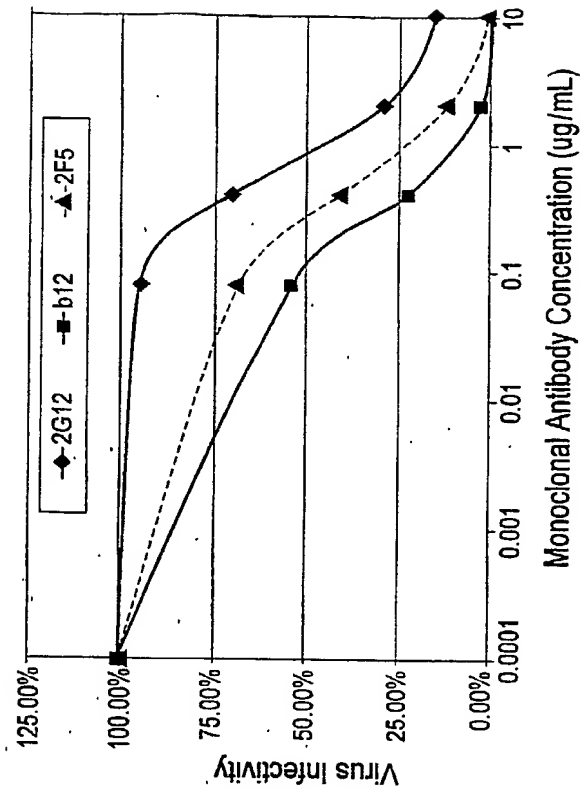
Fig. 23B



Infectivity of virus particles containing the subtype B consensus envelope.

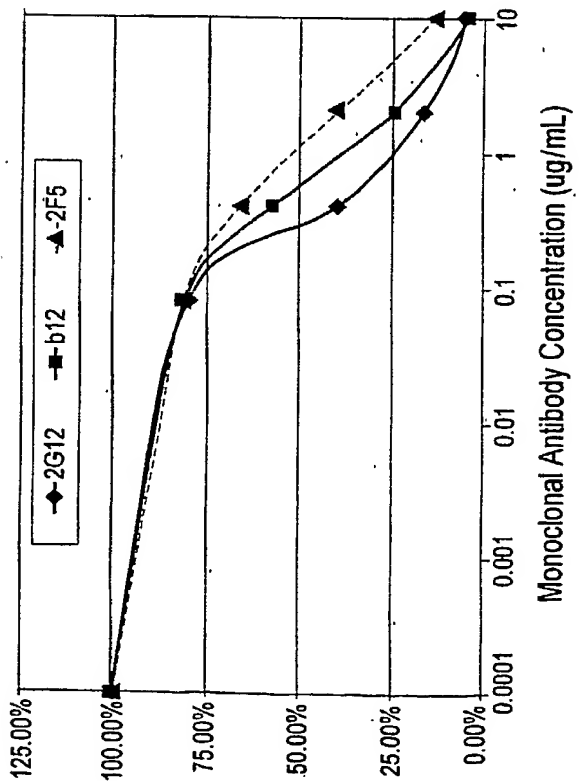
Infectivity of pseudotyped virus containing consensus B gp160 or gp140 was determined using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize  $\beta$ -galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

Fig. 24B



Neutralization of Pseudovirions containing  
NL4.3 Env (gp160)

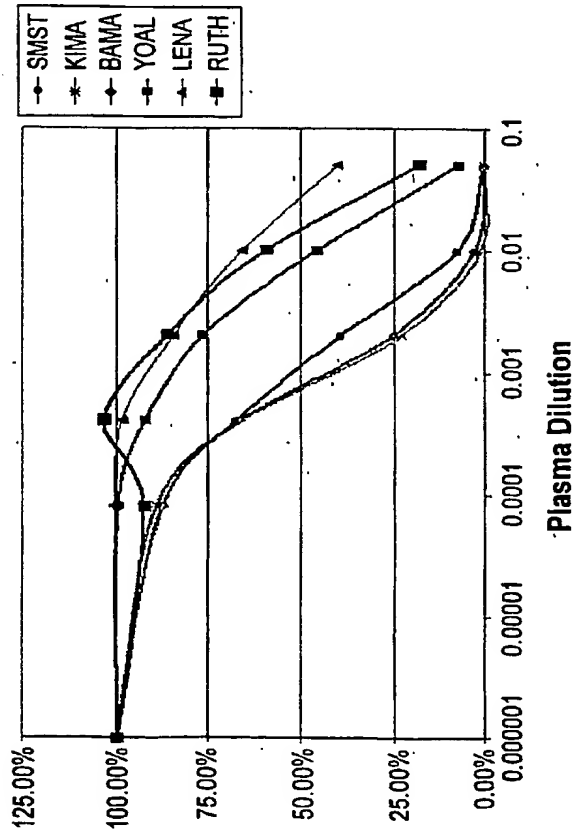
Fig. 24A



Neutralization of Pseudovirions containing Subtype B  
consensus Env (gp160)

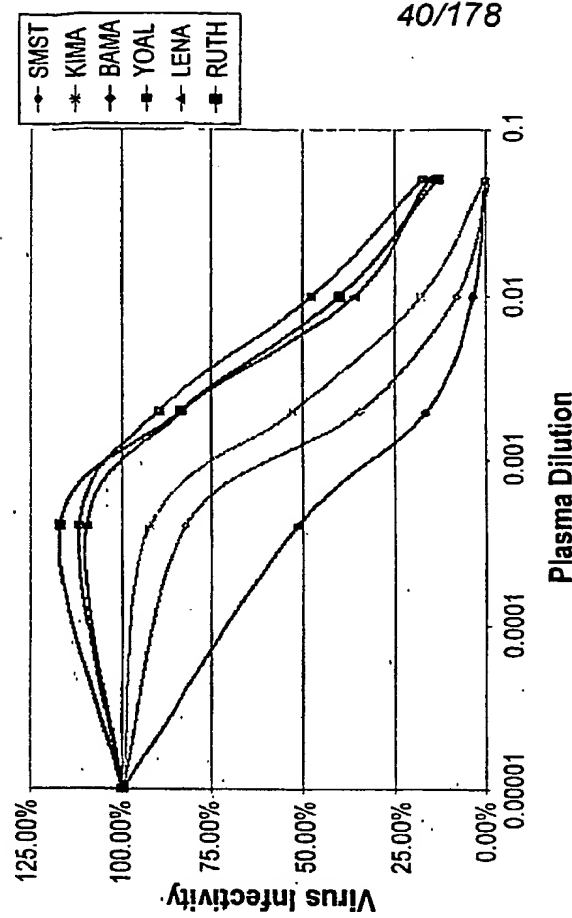


Fig. 24C



Neutralization of Pseudovirions containing Subtype B consensus Env (gp160)

Fig. 24D



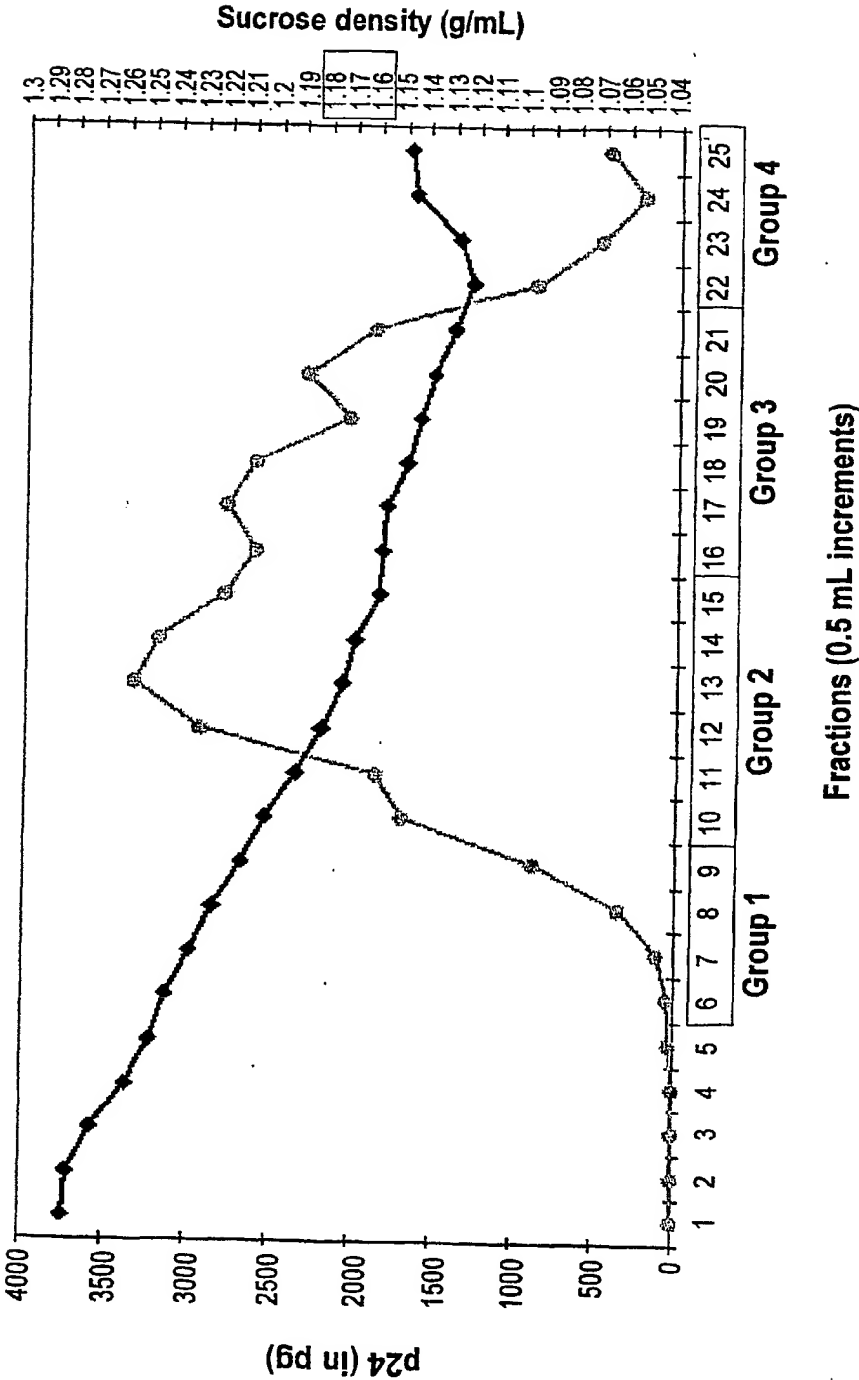
Neutralization of Pseudovirions containing NL4.3 Env (gp160)

#### Neutralization sensitivity of virions containing subtype B consensus gp 160 envelope.

Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1 subtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration ( $IC_{50}$ ) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections.

41/178

Fig. 25A

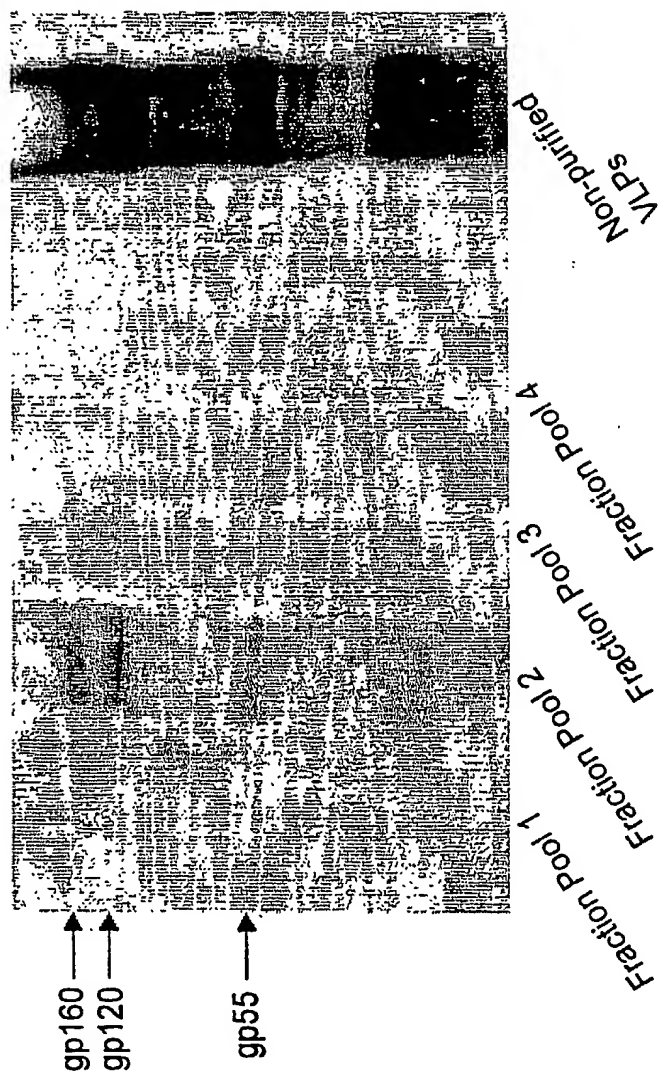


Density and p24 analysis of sucrose gradient fractions.

0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml.

42/178

Fig. 25B



#### VLP production by co-transfection of subtype B consensus gag and env genes.

293T cells were co-transfected with subtype B consensus gag and env genes. Cell supernatants were harvested 48-hours post-transfection, clarified through at 20% sucrose cushion, and further purified through a 20-60% sucrose gradient. Select fractions from the gradient were pooled, added to 20ml of PBS, and centrifuged overnight at 100,000 x g. Resuspended pellets were loaded onto a 4-20% SDS-PAGE gel, proteins were transferred to a PVDF membrane, and probed with plasma from an HIV-1 subtype B infected individual.

43/178

**Fig. 26A****Year 2000 Con-S 140CFI.Env**

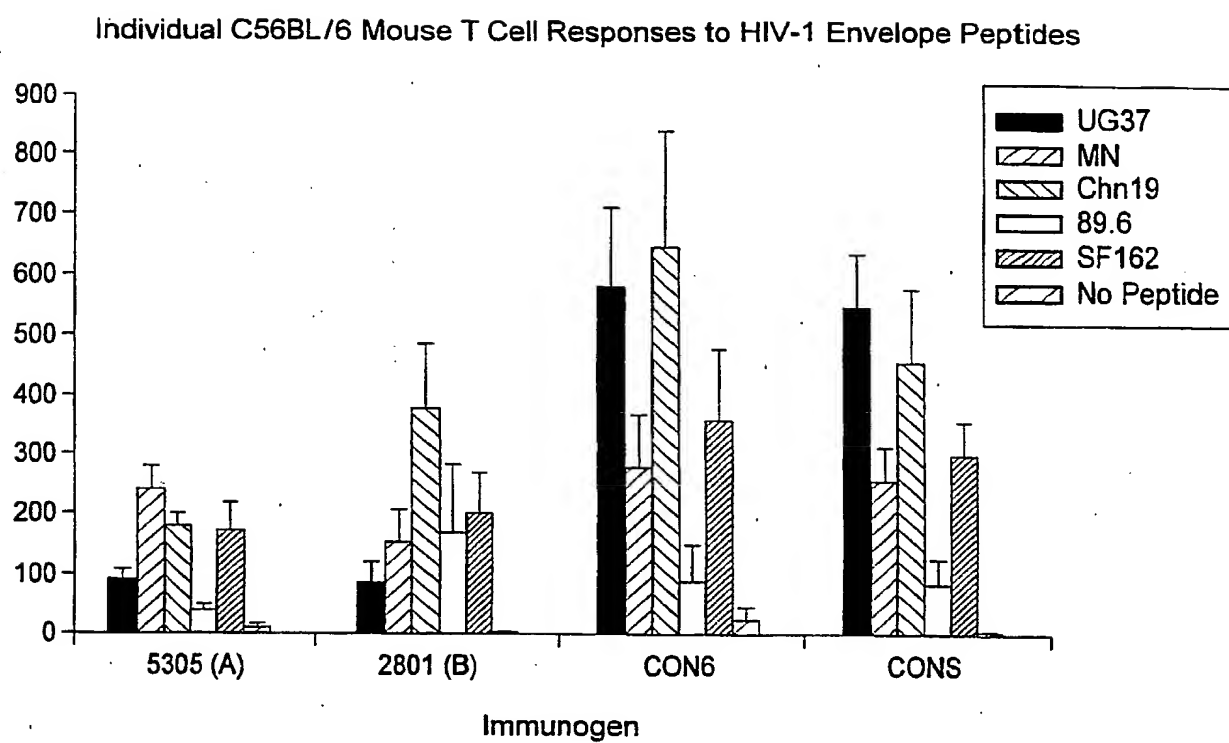
MRVRGIQRNCQHLWRWGTLLGLMLMICSAAENLWVTVYYGVFVWKEANTTLFCASDAKAYDTEVH  
 NVWATHACVPTDPNPQEIIVLENTENFNMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNC  
 TNVNVNTNTNTEEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDNNNNSSNYRLINCNT  
 SAITQACPKVSFEPPIHYCAPAGFAILLKCNDDKFNGTGPCKNVSTVQCTHGKIPVVSTQLLNG  
 SLAEIIIIRSENI TNNAKTIIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQA  
 HCNISGKWNKTLQOVAKKLEHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTW  
 IGNGTKNNNNNTNDTITLPCRIOIINMWQGVQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTN  
 ETEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKLTVQARQLLSGIVQQQSNLLRAIEAQ  
 QHLLQLTVWGKQLQARVLAVERYLKDQQLLEIWDNMTWMEWEREINNYTDIIYSLIEESQNQOEK  
 NEQELLALDKWASLWNWFDITNWLW

A gp140 CFI is referred to HIV-1 envelope design with the cleavage-site-deleted (C), fusion-site-deleted (F) and gp41 immunodominant region-deleted (I) in addition to the deletion of transmembrane and cytoplasmic domains.

**Fig. 26B****Codon-optimized Year 2000 Con-S 140CFI. seq**

ATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCCTGATCCTGGG  
 CATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGT  
 GGAAGGAGGCCAACACCACCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCAC  
 AACGTGTGGGCCACCCACGCCTGCGTGCCACCGACCCCAACCCCAAGGAGATCGTGCTGGAGAA  
 CGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCT  
 CCCTGTGGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCCTGAAGTGC  
 ACCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAAGTGTCTC  
 CTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTCTACCGCCTGG  
 ACGTGGTGGCCATCGACGACAACAACAACACTCCTCAACTACCGCCTGATCAACTGCAACACC  
 TCCGCCATCACCACGAGCCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCACTACTGCGCCCC  
 CGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCTGCAAGAAGC  
 TGTCCACCGTGCAAGTGACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGC  
 TCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAATCACCACCAACGCGCAAGACCATCAT  
 CGTGACAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAAACACCCGCAAGTCCA  
 TCCGCATCGCCCCCGGCCAGGCCTTCTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCC  
 CACTGCAACATCTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGA  
 GCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCC  
 ACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGG  
 ATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCTGCCCTGCCGCATCAA  
 GCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGA  
 TCACCTGCAAGTCCAACATCACCAGGCTGTGCTGACCCGCGACGGCGGCAACAACAACACCAAC  
 GAGACCGAGATCTTCCGCCCCGGCGGCGGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAA  
 GTACAAGGTGGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAAGGCCAAGCTTACCGTGCAGG  
 CCCGCCAGCTGCTGTCCGGCATCGTGACGAGCTGCAACCTGCTGCGCGCCATCGAGGGCCAG  
 CAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCCGTGGA  
 GCGCTACCTGAAGGACAGCAGCTCGAGATCTGGGACAACATGACCTGGAGTGGGAGCGCG  
 AGATCAACAACACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACAGGAGAGAAG  
 AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAAGTGGTTCGACATCACCAA  
 CTGGCTGTGGTGAAGATCC

44/178

*Fig. 27*

45/178

Fig. 28A

Design of expression-optimized HIV-1 envelope gp140CF

Con-B-2003 Env.pep (841 a.a.) \*

MRVKGIRKNYQHLLWRWGTMLLGLMLMICSAAEKLWVTYYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVL  
 ENVTFENFMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTIIYRWRGEIKNCSEFNITTSIRDKVQKEY  
 ALFYKLDVVPIDNDNTSYRLISCNSTVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKFGTGPCTNVSTVQCTHGIRPVVSTQ  
 LLLNGSLAEEEVVIRSENFNDNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGAFYTTGEIIGDIRQAHCNISRAKWNNTLKQ  
 IVKKLREQFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNTWNTTEGNTILPCRIKQIINMWQEVGKAMYAPP  
 IRGQIRCSSNITGLLLTRDGGNNETEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTAKRRVVQREKRAVGIGAMFLGFLGA  
 AGSTMGAASMTLTVQARQLLSGIVQQNNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWCSGKLICTTAVPW  
 NASWSNKSLEIWDNMTWMEWEREIDNYTSLIYTLIEESQNQQEKEQELLELDKWA<sup>SLNWF</sup>DITNWLWYIKIFIMIVGGLVGL  
 RIVFAVLSIVNRVRQGYSPLSFQTRLPAPRGPD<sup>RDPEGIEEGGERDRDRSGRLVDGFLALIWDDLRLSLCLFSYHRLRDL</sup>LLLVTR  
 IVELLGRRGWEVLKYWNLLQYWSQELKNSAVLLNATAIAVAEGTDRVIEVVRACRAILHIPRRIRQGLERALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 28B

Con-B-140CF.pep (632 a.a.)

Nick name: 002

MRVKGIRKNYQHLLWRWGTMLLGLMLMICSAAEKLWVTYYGVVPVWKEATTLFCASDAKAYDTEVHNWATHACVPTDPNPQEVVL  
 ENVTFENFMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTIIYRWRGEIKNCSEFNITTSIRDKVQKEY  
 ALFYKLDVVPIDNDNTSYRLISCNSTVITQACPKVSFEPIPIHYCAPAGFAILKCNCKKFGTGPCTNVSTVQCTHGIRPVVSTQ  
 LLLNGSLAEEEVVIRSENFNDNAKTIIVQLNESVEINCTRPNNTRKSIHIGPGAFYTTGEIIGDIRQAHCNISRAKWNNTLKQ  
 IVKKLREQFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNTWNTTEGNTILPCRIKQIINMWQEVGKAMYAPP  
 IRGQIRCSSNITGLLLTRDGGNNETEIFRPGGDMRDNRSELYKYVVKIEPLGVAPTAK<sup>TLTVQARQLLSGIVQQNNLLRA</sup>  
 IEAQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWCSGKLICTTAVPWNASWSNKSLEIWDNMTWMEWEREIDNYTSLIY  
 TLIEESQNQQEKEQELLELDKWA<sup>SLNWF</sup>DITNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

46/178

## Fig. 28C

Codon-optimized Con-B 140CF.seq (1927 nt.)

Nick name: 002

TTCAGTCGACGGCCACCATGAGGGTGAAGGGTATTCCGGAATAATTACCAACACCTGTGGCGTGGGGAACCATGCTCCTTGGTAT  
 GCTGATGATTGSCAGTCCCGCCGAGAAACTTTGGGTAAGTGTACTACGGCGTTCCTGTCTGGAAGGAAGCTACAACCACTCTT  
 TTTTGTGCATCCGACGCTAAGCTTACGACACAGAAGTGCATAATGTTGGGCCACCCATGCTTGGTCCCTACAGATCCCAACC  
 CCCAGGAAGTCGTCTTGAGAAATGTCAACAGAGAAATTTAAACATGTGGAAGAATAATATGTAACAACAATGCACGAAGACATTAT  
 TAGCCTGTGGGACCACTCTTGAAGCCCTGCGTGAACCTCACTCCACTTTCGCTCACACTTAACCTGTACTGATTGATGAACGCA  
 ACCAACACAAATACATACTATTATATATATCGCTGGAGGGGGAATAACAAGAACTGCTCTTTCAACATCACCACTTCCATAAGGGATA  
 AGGTCCAGAAAGAAATATGCCCTGTTTATAAACTTGATGTGGTCCCGATAGACAATGACAACACTAGCTATCGACTGATCTCTTG  
 TAAACATCCGTGATTACCCCAAGCTTGCCCAAGGTCAGCTTTGAACCAATACCCATTCACTACTGCGCTCCGCTGGTTTGGCC  
 ATCCTCAAGTGTAAACGACAAATAATCAATGGACCGGCTCTCTCGGGAAGAGGAGGTCGTGATTAGAAGCGGAAACTTTACCGGATAACGC  
 CTGTTGTCAGTACCCAACTCTTGAACGGTCTCTCGGGAAGAGGAGGTCGTGATTAGAAGCGGAAACTTTACCGGATAACGC  
 TAAACAATCATTTGTCAACTTAATGAAGCGTCGAAATTAACCTGCACCAAGAGGAGGTCGTGATTAGAAGCGGAAACTTTACCGGATAACGC  
 GGGCCCGCGCGCATTTATACAACTGGCGAAATCACTTGTGACATCAGACAAGCTCATTTGCAATATCTCCGCGGAAATGGA  
 ACAACACCTGAACAGATCGTGAAGAACTTCGAGAACAAATTCGGTAATAAACAATCGTATTCAACCAAGCTCCGGAGGCGA  
 CCTGAGATAGTTATGCACTCATTTCACTGTGGCGCGAGTCTTCTATTGTACACAACCTCAACTTTTAATAGCACTTGAAT  
 GGAACATGGAACAACAAGGGAACATCACTCTGCCCTGTCCGATTAGCAGATCATTAATATGTGGCAAGAAAGTGGGAAAG  
 CTATGTACGCCCCCTATTTCGGGACAAATAAGATGCTCTAGTAATATTACCGGATTGTGCTGACACGCGGAGGAAATAA  
 TGAACAGAGATATTAGACCTGGCGGAGGCGACATGAGAGATAACTGGAGAAAGTGAAGCTTTACAAATATAAAGTCGTAAAGATA  
 GAACCAATGGGGGTAGCACCAACCAAGCAAAACCTTGACAGTACAGGTAGGCAGCTGCTGAGCGGAATCGTGCAACAACAA  
 ATAATCTCTCCGAGCCATAGAACCAACAACATCTGTTGCAGCTGACAGTATGGGGAATCAACAGCTTCAGGCAAGAGTGCT  
 GGCCGTCGAGAGATACCTCAAGATCAACAACCTGCTGGGCATATGGGATGTTCCGGTAAACTCATATGCACACCGCCGTGCC  
 TGGAACGCGAGCTGGTCTAATAATCCCTGGATGAAATTTGGGACAAACATGACTTGGATGGAATGGAACTGGGAAATGACAACT  
 ATACTAGTTTATCTGATCGAAGAACTCTCAGAACCAACAGGAGAAACGAAACAGGAACTGCTGGAACTGGACAAGTG  
 GGCATCATTTGTGGAACCTGTTGACATTACTAACTGGCTGTGGTAAAGATCTTACAA

(For all 140CF design shown here and below, 140CF gene will be flanked with the 5' sequence of "TTCAGTCGACGGCCACC" that contains a Kozak" sequence (GCCACCATGG/A) and SalI site and 3' sequence of TAAAGATCTTACAA containing stop codon and BglII site.)

47/178

Fig. 29A

CON OF CON-S-2003 (829 a.a.)

MRVMGIQRNCQHLWRWGILIFGMLIICSAENLWVTYYGVVWKEANTTLFCASDAKAYDTEVHNWVATHACVPTDPNPQEIVL  
 ENVTFENFMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTNNEEIKNCSEFNITTEIRDKKKVYALFYKL  
 DVVPIDDDNSYRLINCNNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDDKFKNGTGPCKNVSTVQCTHGKIPVSTQLLNGSL  
 AEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCNISRTKWNKTLOQVAKKLRE  
 HENKTIIFNPSSGGDLEITTHSFNCGGEFFCYNTSELFNSTWNGTNNITLPCRRIKQIINMWQGVGQAMYPPIEGKIRCTSNIT  
 GLLLTRDGGNNNTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKARRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITL  
 TVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKQQLGIWGCSCGLICTTNVPWNSSWSNKSQDEI  
 WDNMTWMEWDKEINNNTYTDIISLIEESQNEQELLALDKWASLWNFDTNWLWYIKIFIMIVGGLIGLRIVFAVLISIVNR  
 VRQYSPLSFQTLIPNRPGRDPRPEGIEEGEGEQDRDRSIRLVNGFLALAWDDLRLSLCFSYHRLDLILIAARTVELLGRRGWEA  
 LKYLWNLQYWGQELKNSAISLLDTTIAVAEGTDRVIEVQVRCRAILNIPRRIRQGFERALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 29B

CON-S-2003 140CF.pap (620 a.a.).

Nick name: 006

MRVMGIQRNCQHLWRWGILIFGMLIICSAENLWVTYYGVVWKEANTTLFCASDAKAYDTEVHNWVATHACVPTDPNPQEIVL  
 ENVTFENFMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTNNEEIKNCSEFNITTEIRDKKKVYALFYKL  
 DVVPIDDDNSYRLINCNNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDDKFKNGTGPCKNVSTVQCTHGKIPVSTQLLNGSL  
 AEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHNCNISRTKWNKTLOQVAKKLRE  
 HENKTIIFNPSSGGDLEITTHSFNCGGEFFCYNTSELFNSTWNGTNNITLPCRRIKQIINMWQGVGQAMYPPIEGKIRCTSNIT  
 GLLLTRDGGNNNTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPTAKTLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTV  
 WGIKQLQARVLAVERYLKQQLGIWGCSCGLICTTNVPWNSSWSNKSQDEIWDNMTWMEWDKEINNNTYTDIISLIEESQNEQEK  
 NEQELLALDKWASLWNFDTNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.



48/178

## Fig. 29C

CODON-OPTIMIZED CON-S-2003 140CF.seq (1891 nt

Nick name :006

TTCAGTCGACAGCCACCATGCGGGTCATGGGGATACAGAGGAATTGCCAGCACTTGTGGAGGTGGGGAATTTTGATATTCGGGAT  
 GCTCATATCTGCTCTGCCGCTGAGAACCTGTGGGTCACTGTGTATTACGGCGTTCCCGTCTGGAAAGAGCTAATACTACCCCTG  
 TTTTGTGCAAGCGACGCCAAAGCATACGACACCGAAGTCCACAATGTCTGGGTACCCACGCTGTGTACCTACTGATCCAAATC  
 CCCAGGAAATTGTTCTTGAAAACGTAAACGGAACCTTAAACATGTGGAAGAATAATATGGTGGAGCAAAATGCACGAGGATATAAT  
 CAGCCTGTGGGACCAAGTCCCTCAAAACCATGCGTTAAACTCACTCCACTGCGTGACTCTGAACCTGTACCGACGTGAACGCAACC  
 AATAATACAACAACAATGAGGAGATAAGAAATTGTTCAATTAATAATAACCACTGAGATACGGGATACGAAATAAGGTTTATG  
 CACTCTTTTACAAGCTCGACGTGGTGCCCATAGACGACAATAATAGCTACCGACTCATTAATTGCAATACTAGCGCTATAACCCA  
 GGCATGCCCCAAAGTTTCTTCGAGCCCATACCGATTCACTACTGCGCACCCGCGGATTGCGCATTTCTTAAATGCAATGACAAG  
 AAGTTCAACGGCACCGAACCTGTAAAGAACGTAAAGCACTGTTCAATGTACACATGGAATTAAGCCGGTAGTGTCAACGCACTCC  
 TCCTCAACGGGAGCCTTGCAAGAAGAAGAGATCATATCAGGTCAGAAATAATCACTAACACGCGAAACAATCATTTGTTCAAGCT  
 GAATGAGTCTGTAGAAATCAATTGTACCCGCCCTAATAATAACACAAGAAAGTCAATTAGGATCGGACCCGCGGAGCTTCTAC  
 GCAACCGGAGATATCATCGGGGATATACGACAGGCCCACTGCAACATTTCTAGAACTAAGTGAATAAACTTTGCAGCAGTAG  
 CCAAGAAACTGCGGGAACATTTTAAATAAGACAATCATCTTCAATCCAAGTAGCGGAGGGACCTGGAAATCACTACACATTCCTT  
 TAACTGTGGGGCGAGTTTCTACTGTAAATACCTCTGAACCTTCAACATGGAATGGCACTAACCAATACTATAAATCTT  
 CCTTGCAGAAATAAACAGATTATCAACATGTGGCAGGGTGTGGGCAAGCAATGTATGCACCACTCAATCGAAGGCAAAATAAGAT  
 GCACCTCCAATATTACCGGACTCCTCTGACACGGGATGGCGAAACAATAACACGGAGACCTTTAGGCCAGGCGGCGGATAT  
 GAGAGATAACTGGCGCTCCGAGCTCTATAAATAACAAGTCGTTAAGATCGAGCCCTTGGAGTTGCGCAACCAAGCTAAAACC  
 TTGACCGTGCAAGCCAGGCAAGTTGTTGTACAGGTATCGTACAGCAAGCAATCTAATCTTTTGAAGCCATTGAGGCTCAGCAGCAC  
 TCTTGCAGCTTACCGTCTGGGGCATCAACAACCTTACGGCACGCTCCTGGCGGTAGAGCGCTATTGAAAGACCAACAATCTCT  
 CGGGATCTGGGGTGTCTGGAAAATTGATCTGCACGACAAATGTGCCTTGGAACACAGCAGCTGGTCAAAATAAAGCCAAAGACGAA  
 ATATGGGATAACATGACATGGATGGGATAAAGAAATTAATAATTACACTGACATTAATTTACTCACTTATCGAGGAATCAC  
 AAAATCAACAGGAAAAAAATGAACAGGAACTCTTGGCTCTGGACAATGGGCTTCACCTGTGGAACCTGGTTCGACATCACAAATTG  
 GCTCTGGTAAAGATCTTACAA



50/178

## Fig. 30C

CODON-OPTIMIZED Con-A1-2003.seq

Nick name: 001 (1918 nt)

TTCCAGTCGACAGCCACCATGAGGTGATGGGAATCCAACGGAACGCCAGCATCTTCTCCGGTGGGAAACGATGATACCTGGGAAT  
GATAATAATCTGCTCTGCCGCTGAAACCTCTGGGTACAGTGTACTACGAGTGCCTGTATGGAAGGACGCTGAAACCACTCTC  
TTTGTGCTTCCGATGCTAAAGCTACGAAACCGAGATGCACAAATGTTTGGGCCACCCACGCTGCGTGCCAACTGATCCTAATC  
CACAAAGAAATACATCTGGAGAATGTTACTGAGGAATTTAATGAGGAAATTAATGAGGAAATGCACACTGACATCAT  
TTCACTCTGGGACCAATCACTCAAAACCCCTGCTGCTTAACTTACCCCTCTGCGTGACCTCAATTTGTAGCAACGTCAACGTCACA  
AATAATACAACCAACACTCACGAGGAAGAAATTAATAATGCTCTTAAATATGACCACTGAACCTTCGGGACAAATAAACAAG  
TCTATTCACTGTTTTATAGGCTGGACGTCGTCCAAATCAACGAGAACAAATCTAACAGTAGCTATCGACTTATCAATTGCAATAC  
CTCTGCTATTACCCAGCTTGTCTAAAGTCTCTTTGAACCAATCCCTATCCACTACTGTGCCCCAGCTGGATTTCGCAATTCTG  
AAGTCAAGGATAAGGAATTCACGGAACCTGGCCCTTGGCAAGAACGTTAGCACTGTCCAATGCACTACGGAATCAAAACCACTAG  
TCAGCACTCAACTGCTCCTGAATGGCTCACTCGCCGAAGAGAGGTGATTATCCGAAGCGGAGAACATAAATAACAAATGCGAAGAC  
AATAATTGTTCAATTGACGAACCCAGTGAAGATCACTGTACTAGACCAATAACAACAAGAAATCTATCAGAATTGGCCCC  
GGACAAGCCTTCTACGCAACAGGAGATATCATAGGTGACATCAGACAGGCCCATTTCAAGAGCGAGTGGAAATAAA  
CACTCCAGAAAGTGGCAAGCAGCTGAGAAATACTTTAAGAAACAAGACAATCATATTTACTAACTCCTCCGGAGTGATCTCGA  
AATAACCACTCATAGCTTTAATTGTGGGGCGAATCTTCTACTGTAACACATCTGGCCCTCTTTAATTCTACCTGGAATAACGGC  
ACCATGAAAATACTATCACCTCCCTTGCAGAAATTAAGCAAAATCATTAACATGTGGCAGAGCAGGACAGGCCCATGTATGCC  
CTCCCATTTCAAGGTGTGATTTCGATGTGAAGCAACATTAAGCACTTCTTCTGACCCGGGATGGCGAAATAATAATACCAATGA  
GACATTCAGACCCCGCGCGGATATGCGAGACAAATGGCGAAAGTGAACCTTTATAAATAACAAGTAGTTAAGATTGAGCCCCCTT  
GGAGTGGCCCTACTAGAGCAAAAACATTGACCGTTTCAGGCCAGGCGAGTGTCTCAGGAATCGTGACGCAAGTAACCTCC  
TCCGAGCTATCGAGGCACAACAACATCTCTTGAATTTGACCGTATGGGGAATCAAGCAATTCAGGCTAGGGTTTGGCTGTGA  
ACGCTATCTCAAGGATCAGCAGCTTCTGGGAATCTGGGATGCTCTGGGAAATTTGAGAAATTTGATGTAACAAACCTGGAACCTCA  
AGCTGGAGTAATAAAGCCAGAACGAAATTTGGGATAATATGACCTGGCTGCAGTGGGACAAAGAAATTTCTAATTATACTCATA  
TCATATACAATCTGATCGAAGAAATCACAGAACCCAGGAAAGAAATGAGCAAGACCTTCTGGCCCTTGGACAAGTGGGCTAACTT  
GTGGAACCTGGTTGACATTAGCAACTGGCTGTGGTAAAGATCTTACAA

51/178

Fig. 31A

**CONSENSUS C-2003 (835 a.a)**

MRVRGILRN<sup>C</sup>QQW<sup>I</sup>WILGFWMLMCNVVGNLWTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNWVWATHACVPTDPNPQEIVL  
 ENVTFENFMWKN<sup>D</sup>MVDQM<sup>H</sup>EDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEIKNCSEFNITTEL<sup>R</sup>DKKQKVYALFYRLDI  
 VPLNENNSYRLIN<sup>C</sup>NTSAITQACP<sup>K</sup>VSEFDP<sup>I</sup>PIHYCAPAGYAILKCN<sup>N</sup>KTFNGTGPCNNVSTVQCTHG<sup>I</sup>KPVVSTQLLNGSLAE  
 EEIIIRSENLTNNAK<sup>T</sup>IIVHLNESVEIVCTRPNNNTRK<sup>S</sup>IRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQK<sup>V</sup>SKKLKEHF  
 PNKTIKFEPSSGGDL<sup>E</sup>ITTHSFNCRGEFFYCNTSKLFNSTYNST<sup>I</sup>TLP<sup>C</sup>RIKQIINMWQEVGRAMYAPPIAGNITCKSNITG  
 LLLTRDGGKNN<sup>T</sup>ETFRPGGDMRDNRSELYKYK<sup>V</sup>VEIKPLGIAPTAKRRVVEREKRAVGIGAVFLGELGAAGSTMGAASITLT  
 VQARQLLSGIVQQSNLLRAIEAQ<sup>Q</sup>HMLQ<sup>L</sup>TVWGIKQLQTRVLAIERYLK<sup>D</sup>OQLLGIWCGSGKLICTTAVPWNSSWSNKSQ<sup>E</sup>DIW  
 DNMTWMQWDREISNYTDTIYRLLED<sup>S</sup>Q<sup>Q</sup>QEKNEKDL<sup>L</sup>ALDSWKNLWNWFDITNWLWYIKIFIMIVGG<sup>L</sup>IGLRIIFAVLSIVNRV  
 RQYSP<sup>L</sup>SFQTLTPNPRGPDRLGRIE<sup>E</sup>EGGEQDRDRSIRLVSGFLALAWDDLRSLC<sup>L</sup>FSYHRLRDFILIAARAVELLGRSSLRGL  
 QRGWEALKYLGSLVQVWGLELKKSAISL<sup>D</sup>TIAIAVAEGTDRIELIQRICRAIRNIPRRIRQGF<sup>E</sup>AALQ

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design..

Fig. 31B

**Con-C 2003 140CF.pap (619 a.a.)****Nick name: 003**

MRVRGILRN<sup>C</sup>QQW<sup>I</sup>WILGFWMLMCNVVGNLWTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNWVWATHACVPTDPNPQEIVL  
 ENVTFENFMWKN<sup>D</sup>MVDQM<sup>H</sup>EDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEIKNCSEFNITTEL<sup>R</sup>DKKQKVYALFYRLDI  
 VPLNENNSYRLIN<sup>C</sup>NTSAITQACP<sup>K</sup>VSEFDP<sup>I</sup>PIHYCAPAGYAILKCN<sup>N</sup>KTFNGTGPCNNVSTVQCTHG<sup>I</sup>KPVVSTQLLNGSLAE  
 EEIIIRSENLTNNAK<sup>T</sup>IIVHLNESVEIVCTRPNNNTRK<sup>S</sup>IRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQK<sup>V</sup>SKKLKEHF  
 PNKTIKFEPSSGGDL<sup>E</sup>ITTHSFNCRGEFFYCNTSKLFNSTYNST<sup>I</sup>TLP<sup>C</sup>RIKQIINMWQEVGRAMYAPPIAGNITCKSNITG  
 LLLTRDGGKNN<sup>T</sup>ETFRPGGDMRDNRSELYKYK<sup>V</sup>VEIKPLGIAPTAK<sup>T</sup>LTVQARQLLSGIVQQSNLLRAIEAQ<sup>Q</sup>HMLQ<sup>L</sup>TVW  
 GIKQLQTRVLAIERYLK<sup>D</sup>OQLLGIWCGSGKLICTTAVPWNSSWSNKSQ<sup>E</sup>DIWDNMTWMQWDREISNYTDTIYRLLED<sup>S</sup>Q<sup>Q</sup>QEK<sup>N</sup>  
 EKDL<sup>L</sup>ALDSWKNLWNWFDITNWL<sup>W</sup>\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

52/178

## Fig. 31C

CODON-OPTIMIZED Con-C-2003 140CF (1,888 nt.)

Nick name:003

TTCTAGTCGACAGCCACCATGCGAGTGAGAGGCCATTCTGCGGAATTGTCAGCAATGGTGGATCTGGGGCATACTCGGATTCTGGAT  
GCTTATGATATGCAATGTTGTGGGAACCTGTGGGTACCGTATATCTAGTCTGGGTTCCAGTCTGGAAGGAGGCTAAACAACACGCTG  
TTCTGTGCAAGTGACGCCCAAGCCTACGAGAAAGAGTGCACAACGCTGTGGGTACCCACGCTTGTGTTCGAACCGATCCAAACC  
CCAGGAAATCGTCTCGAGAACGTGACTGAAACCTTTAACATGTGGAAGATGATATGGTAGATCAGATGCACGAAGATATCAT  
TTCAATTGTGGACCAATCATTTGAAACCATGCGTAAACTGACCCCTCTGCGTAACACTTAAGTGCACCAATGCAACTTAATGCC  
ACCAATACTATGGCGGAAATAAAAACTGTAGCTTTAACATTACAACGGAATCCGGGATAGAAACAAAGGTCTACGCGCTCT  
TTTACCGACTCGATATCGTCCACTTAACGAGATAATAGTTACCGCTGTATTAAGTGTAAACACATCAGCCATTACGCAAGCTTG  
CCCCAAGTTTCTTCGACCCCATCCCAATTCACTATTGTGCCCCCGCTGGATACGCTATACTTAAATGCAACAATAAAACATTT  
AATGGAACCGGACCATGTAAACAACGTCAGTACCGTACAATGTACGACGGAATTAAACCTGTGTCTCAACCCAGCTTCTCCTTA  
ACGGCTCATTTGGCGGAGGAAGAAATTATTATCAGATCAGAAACCTTGACCAACAATGCCAAAACCATCATCGTGCACTCAATGA  
ATCCGTGGAATCGTGTGACCGAGACCAATAACAAATACCGGAATCAATCAGGATTTGGGCTGGCCAGACATTTACGCTACA  
GGTGATATAATTGGCGGATATTAGACAAAGCCCATTTGCAACATATCAGAAAGACAAGTGAATAAGACTCTGCAGAAAGTTTCTAAGA  
AGCTGAAGGAACACTTTCCCAATAAAACGATTAAAGTTGAGCCCTCTTCAGGAGGAGACCTTGAGATCAACAACACTCTTTTAA  
TTGTAGAGGGGAGTTCTTCTATTGTAATACATCAAGCTCTTAAACAGTACCTACAACCTCCACTAATAGTACCATCACACTCCC  
TGCAGATAAAGCAATAATCAACATGTGGCAAGAGTTGGCCGAGCAATGTAGCCCCCTCCCATCGGAGCAACATTACATGTA  
AATCCAATATTACTGGCCTTTTGTGACACGGGACGGGAAAGATAACACTGAGACCTTCAGACCTGGCGGAGCGGATATGCG  
CGATAATTGGCGGAGAGCTCTACAAGTATAAAGTCGTTGAAATCAAGCCACTGGGCATAGCTCCTACGAAAGCAAGACACTC  
ACTGTTCAGGCTAGACAGCTGCTCTCCGGCATAGTGCAACAGCAATCCAATCTCCTGCGAGCTATCGAAGCCCAACATATGC  
TCCAGCTTACCGTCTGGGGAATCAACAATTCGAACACGAGTGTGGCGGATAGAGAGATATTTGAAAGATCAGCAACTCCTGGG  
GATTTGGGGCTGTTTCAGGTAAGCTCATCTGTACAACCTGCGGTGAACTCAAGCTGGAGTAAACAAGCAAGAGGATATA  
TGGGACAACATGACTTGGATCGAGTGGGATCGAGAAATAAGCAACTATACAGATACCATTTATCGGCTCCTGGAGGACTCACAGA  
ACCAGCAGGAGAAAAATGAGAAAGATTGTCTCGCGCTTGACAGTTGGAAGAAATTTGTGGAATTGGTTCGACATTACAAACTGGCT  
CTGGTAAAGATCTTACAA

53/178

Fig. 32A

**CONSENSUS G-2003 (842 a.a.)**

MRVKGIO<sup>W</sup>QH<sup>L</sup>LK<sup>W</sup>GT<sup>L</sup>LILGLV<sup>I</sup>ICSASNNLW<sup>T</sup>VYGV<sup>P</sup>WEDAD<sup>T</sup>TLFCASDAKAYSTERHN<sup>V</sup>WATHACV<sup>P</sup>TD<sup>P</sup>NPQEIT<sup>L</sup>  
 ENV<sup>T</sup>ENFNM<sup>W</sup>KNNM<sup>V</sup>EQMHEDIISLWDESLK<sup>P</sup>CVK<sup>L</sup>TLPLCV<sup>T</sup>LNCTDV<sup>N</sup>TNNNTN<sup>T</sup>NTKKEIKNCSE<sup>N</sup>ITTEIR<sup>D</sup>DKKKKEYALFY  
 RLDV<sup>P</sup>INDNGNSSIYRLIN<sup>C</sup>NVSTIKQACP<sup>K</sup>VTFDPIPIHYCAPAGFAILKCRD<sup>K</sup>KFNGTGPC<sup>N</sup>KNVSTVQCTHGI<sup>K</sup>PPVSTQ<sup>L</sup>L  
 L<sup>NG</sup>SLAEE<sup>E</sup>IIIRSENITD<sup>N</sup>TKV<sup>I</sup>IVQLNETIEIN<sup>C</sup>TRPNN<sup>T</sup>RK<sup>S</sup>IRIGPGQAFYATGDIIGDIRQAHCN<sup>V</sup>SR<sup>T</sup>KNWEM<sup>L</sup>QK<sup>V</sup>K  
 AQLK<sup>K</sup>IFNK<sup>S</sup>ITFNSSGGDL<sup>E</sup>ITTHSFNCRGEFFYC<sup>N</sup>TSGLFNN<sup>S</sup>LLNST<sup>T</sup>ITLPC<sup>K</sup>IKQIVRM<sup>W</sup>QRVGQAMYAP<sup>P</sup>PIAGN<sup>I</sup>T  
 CRSNITGL<sup>L</sup>LTRDGGNN<sup>T</sup>ETFRPGGDM<sup>R</sup>DNWRSELYK<sup>I</sup>KV<sup>I</sup>KIP<sup>L</sup>GVAP<sup>T</sup>RARR<sup>R</sup>VEREKRAVGLGAV<sup>L</sup>LGLGAAGST<sup>M</sup>G  
 AASIT<sup>L</sup>TVQVRL<sup>L</sup>SGIVQQ<sup>S</sup>NLLRAIEAQH<sup>L</sup>LQ<sup>L</sup>TVG<sup>I</sup>KQ<sup>L</sup>QARVLAVERYLKDQ<sup>L</sup>LGIW<sup>G</sup>CSGK<sup>L</sup>ICTTN<sup>V</sup>PWNT<sup>S</sup>WSN  
KS<sup>Y</sup>NEIWDNMTWIEWEREISNYTQ<sup>I</sup>YSLIEESQ<sup>N</sup>QEKNEQ<sup>D</sup>LLALDKWASLWNWF<sup>D</sup>ITKWLW<sup>I</sup>KIFIMIVGGLIGLRIVFAV  
LSIVNRVRQGYSP<sup>L</sup>SFQ<sup>T</sup>LTHHQREPD<sup>R</sup>PERIEEGGEGQ<sup>D</sup>KDRSIRLVSGFLALAWDDLRS<sup>L</sup>CLFSYHRLRDFILIAARTVELLG  
RSSLKGLRLGWGLKYLWNL<sup>L</sup>LYWGQELKNSAINLLDTIAIAVANWTD<sup>R</sup>VIEVAQRA<sup>C</sup>RAILNIPRRIRQGLERALL  
 \*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF  
 design and the "W" underlined with red color is the last amino acid at the C  
 terminus, and all the remaining amino acids after the "W" will be deleted in 140CF  
 design.

Fig. 32B

**Con-G-2003 140CF (626 a.a.)****Nick name: 007**

MRVKGIO<sup>W</sup>QH<sup>L</sup>LK<sup>W</sup>GT<sup>L</sup>LILGLV<sup>I</sup>ICSASNNLW<sup>T</sup>VYGV<sup>P</sup>WEDAD<sup>T</sup>TLFCASDAKAYSTERHN<sup>V</sup>WATHACV<sup>P</sup>TD<sup>P</sup>NPQEIT<sup>L</sup>  
 ENV<sup>T</sup>ENFNM<sup>W</sup>KNNM<sup>V</sup>EQMHEDIISLWDESLK<sup>P</sup>CVK<sup>L</sup>TLPLCV<sup>T</sup>LNCTDV<sup>N</sup>TNNNTN<sup>T</sup>NTKKEIKNCSE<sup>N</sup>ITTEIR<sup>D</sup>DKKKKEYALFY  
 RLDV<sup>P</sup>INDNGNSSIYRLIN<sup>C</sup>NVSTIKQACP<sup>K</sup>VTFDPIPIHYCAPAGFAILKCRD<sup>K</sup>KFNGTGPC<sup>N</sup>KNVSTVQCTHGI<sup>K</sup>PPVSTQ<sup>L</sup>L  
 L<sup>NG</sup>SLAEE<sup>E</sup>IIIRSENITD<sup>N</sup>TKV<sup>I</sup>IVQLNETIEIN<sup>C</sup>TRPNN<sup>T</sup>RK<sup>S</sup>IRIGPGQAFYATGDIIGDIRQAHCN<sup>V</sup>SR<sup>T</sup>KNWEM<sup>L</sup>QK<sup>V</sup>K  
 AQLK<sup>K</sup>IFNK<sup>S</sup>ITFNSSGGDL<sup>E</sup>ITTHSFNCRGEFFYC<sup>N</sup>TSGLFNN<sup>S</sup>LLNST<sup>T</sup>ITLPC<sup>K</sup>IKQIVRM<sup>W</sup>QRVGQAMYAP<sup>P</sup>PIAGN<sup>I</sup>T  
 CRSNITGL<sup>L</sup>LTRDGGNN<sup>T</sup>ETFRPGGDM<sup>R</sup>DNWRSELYK<sup>I</sup>KV<sup>I</sup>KIP<sup>L</sup>GVAP<sup>T</sup>RART<sup>L</sup>TVQVRL<sup>L</sup>SGIVQQ<sup>S</sup>NLLRAIEAQ<sup>QH</sup>  
LLQLTVG<sup>I</sup>KQ<sup>L</sup>QARVLAVERYLKDQ<sup>L</sup>LGIW<sup>G</sup>CSGK<sup>L</sup>ICTTN<sup>V</sup>PWNT<sup>S</sup>WSNKS<sup>Y</sup>NEIWDNMTWIEWEREISNYTQ<sup>I</sup>YSLIEES  
Q<sup>N</sup>Q<sup>N</sup>Q<sup>E</sup>KNEQ<sup>D</sup>LLALDKWASLWNWF<sup>D</sup>ITKWLW<sup>I</sup>\*

\*Amino acids seen in blue color is for easy identification of the junction of the

deleted fusion cleavage site

54/178

## Fig. 32C

CODON-OPTIMIZED Con-G-2003 140CF.seq

Nick name:007

TTCAGTCGACAGCCACCATGCGAGTGAAGGGAATCCAGAGAAATTGGCAGCACCTTTGGAAAGTGGGGCACACTCATCTCGGCCT  
TGTGATCATATGCTCTGCCTCAAAATAACCTTTGGGTACAGTTTATTACCGCGGTGCCCTTTGGAGGACGAGACACAACACTCTT  
TTTTGTGCCAGCGACGCTAAGGCTTATTCAACAGAGAGGCATAACGTTTGGGTACACATGCGTCCGACCGATCCTAATC  
CCAGGAAATCACTCTTGAGAAATGTTACAGAGAATTTAATATGTGGAAGAACACATGGTTGAACAGATGCATGAAGACATAAT  
TTCTCTCTGGGATGAATCTCTGAAACCTTGCGTGAAGCTTACACCACTGTGCGTTACCTGAAATGCACTGACGTCAATGTCA  
AATAATAATACCAACAATACAAAAAAGAAATCAAAATTTGTTCTTCAACATAACCAACCGAGATACCGATAAAAAAAGAAAG  
AATACGCCCTGTTCTACAGACTCGATGTGTCCTCCCAATTAATGACAAACGGAATTTCTTCCATCTACCGACTTATCAATTGTAACGT  
GTCTACAATCAACAGGCCTGTCCCTAAAGTCACATTTGACCCCTATTCCCATTTCAATTAATGCTGCCCCCGCTGGCTTCTT  
AATGCCGAGACAAAAAATTTAACGGAACAGGACCATGCAAGAATGCTCAACAGTTCAATGCACTCATGGAATTAACACGTCG  
TTTCTACTCAACTCCTTCTCAATGGAAGCCTGGCAGAAGAGGAATCATATCCGAGCGAAACATAACCGACAAACACAAAGT  
AATCATCGTACAGCTGAACGAGACCATTTGAATAAATTTGTACGAGACCTAATAATAACAAGAAAAAGCATACGCATCGGCCCC  
GGACAGGCTTTCTACGCCACAGGAGACATTATCGGAGATATCGCCAGGCTCACTGTAATGTCTAGAACAAAAATGGAACGAAA  
TGCTTCAGAAAGTCAAAGCTCAGCTCAAGAAATATTCAACAAATCTATTACATCAACTCATCATCAGGCGGCGATCTGGAGAT  
AACAACTCATTCCTTCAACTGTCGGGAGAAATTTTTTACTGTAACACGTCGCGCTGTTCAACAATTCACCTCCTGAATAGCACT  
AACTCCACCATCACTCTCCCATGTAAGATCAACAAATCGTCAGAATGTGGCAGCGAGTCGGTCAAGCTATGTACGCCCCTCCAA  
TCGCCGGTAATATCACATGTAGAAGCAATATCACAGGCTCTTGCTTACAAGGACGCGGGAAACAACAACCCGAAACCTTCAG  
ACGAGGAGGAGACATGCGAGACAAATTTGGCGGAGCGAGCTGTATAAATATAAGATCGTAAATCAAACTTGGGTGAGCG  
CCAACTAGAGCCCGAACACTGACCGTGAGGTGAGGCAACTGCTGAGCGGCTGTCCAAACAACATCCAACTCTTAGAGCAA  
TCGAGGCCAGCAGCATCTGCTCCAGCTTACTGTATGGGAATCAACAACACTGCAAGCAAGAGTATTGGCAGTGGAGAGGTATCT  
CAAGGACCAGCAGCTTCTGGGAATTTGGGTTGCAGCGGAAAGCTCATATGTACAACCAATGTGCCCTGGAACTAGTTGGAGT  
AATAAGAGTTACAATGAATCTGGACAATATGACATGGATCGAATGGAGCGGAAATATCCAACCTATCTCAGCAAAATCTATT  
CCCTCATTTGAAGAGAGTCAGAACCCAGCAGGAAAAAGAAATGAGCAAGACCTTCTCGCCCTGGATAAAATGGGCATCTCTGTGGAACCTG  
GTTTGACATAACTAAATGGTTGTGGTAAAGATCTTACAA

55/178

Fig. 33A

CONSENSUS 01 AE-2003 (854 a.a.)

MRVKETQMNWPNLWKWGTLILGLVICSASDNLWVTYYGVVWRDADTTLCASDAKAHETEVENNVWATHACVPTDPNPQEIHL  
 ENVTFENFMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTNNNITNVSNIIIGNITNEVRNCSFNMTTTELDRDKK  
 QKVHALFYKLDIVQIEDNNSYRLINCVTSVIKQACPKEISFDPIPIHYCTPAGYAILKCNCKNFNGTGPCKNVSSVQCTHGKIPVV  
 STQLLNGSLAEEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEV  
 LKQVTEKLKEHFNNKTIIFQPPSGGDLEITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIIILPCKIKQIINMWQAGQA  
 MYAPPISGRINCVSNITGILLTRDGGANNETFRPGGNIKDNWRSELYKYKVQIEPLGIAPTRAKRRVVEREKRAVGIGAMI  
 FGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKDKQFLGLWGCSGKIIC  
 TTAVPWNSTWSNRSEIWNMTWIEWEREISNYTNIQIYEILTESQOQDRNEKDLLELDKWASLWNNWFDITNWLWYIKIFIMIV  
 GGLIGLRIIFAVLSIVNRVQGYSPLSFQTPTHHQREPDPRPERIEEGGEQGRDRSVRLVSGFLALAWDDLRLSLCLFSYHRLRDF  
 ILIAARTVELLGHSSSLKGLRRGWEGLKYLGILLNLLYWGQELKISAILLATAIAVAGWTDRIEVAQQAARAILHIPRRIRQGLE  
 RALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted as 140CF.

Fig. 33B

Con-AE01-2003 140CF.pap (638 a.a.)

Nick name: 008

MRVKETQMNWPNLWKWGTLILGLVICSASDNLWVTYYGVVWRDADTTLCASDAKAHETEVENNVWATHACVPTDPNPQEIHL  
 ENVTFENFMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTNNNITNVSNIIIGNITNEVRNCSFNMTTTELDRDKK  
 QKVHALFYKLDIVQIEDNNSYRLINCVTSVIKQACPKEISFDPIPIHYCTPAGYAILKCNCKNFNGTGPCKNVSSVQCTHGKIPVV  
 STQLLNGSLAEEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEV  
 LKQVTEKLKEHFNNKTIIFQPPSGGDLEITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIIILPCKIKQIINMWQAGQA  
 MYAPPISGRINCVSNITGILLTRDGGANNETFRPGGNIKDNWRSELYKYKVQIEPLGIAPTRAKTLTVQARQLLSGIVQQQ  
 SNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKDKQFLGLWGCSGKIICCTTAVPWNSTWSNRSEIWNMTWIEWEREISN  
 YTNQIYEILTESQOQDRNEKDLLELDKWASLWNNWFDITNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.



56/178

## Fig. 33C

CODON-OPTIMIZED Con-AE01-2003 140CF.seq (1945 nt.)

Nick name: 008

ttcagtcgacagccaccatgCGAGTCAAGGAAACACAAATGAACCTGGCCTAATCTGTGAAGTGGGCAACCCCTGATCCTGGGTTT  
GGTCATTATTTGCTCTGCGAGCGACAATCTCTGGTTACTGTCTATTACGGAGTCCCGTTTGAGAGATGCCGACACTACACTG  
TTCTGCGCTCAGATGCCAAAGCTCATGAACCTGAAGTGCATAATGTTTGGGCAACCCACGCTGTGTCTTACCGACCCCAACC  
CCCAAGAAATACACCTTGGAACCGTGACCGAGAACTTTAATATGTGAAGAATAACAATGTTGAACAGATGCAAGAAGACGTAAT  
CAGCCTGTGGGATCAAGTCTGAACCTTGCGTAAACCTGACTCCACTTTCGTAACACTTAATTCACCAACCGGAACCTGACA  
AACGTTAACACATCACTAACGTTCCACATCATCGGCAACATAACGAACGAAGTGAGAAATTCAGTTCATATGACTACAG  
AGCTCCGGGACAAGAACAGAGGTCCTCTCTTTTACAACTCGACATCGTCCAGATCGAAGACAATAACAGCTACAGACT  
TATAAATTGTAATACATCCGTGATTAAACAAGCATGCCCCAAAATAAGCTTCGATCCTATTCTATCCACTACTGTCTGCCC  
GGCTATGCTATCTTGAAATGCAATGATAAGAACTTCAATGGGACCGGACCTTGTAAAGACGTGTCTAGTGTGCAATGCACTCAGG  
GCATTAAACCAAGTGGTAAGCACCCAGCTGCTCTGGAACGGCTCTGCGAGAGGAGATTAATTCGAAAGTGAGAAACCTCAC  
CAACAACGCTAAGACTATCATCGTACATCTCAATAAATCAGTCGAAATTAATTCACCCAGACCCCTCCAATAATACTAGAACTTCA  
ATCACTATCGGCCAGGACAAGTCTTTATAGAACAGGAGATATCATAGGAGATATCAGAAAGGCATATTGCGAGATAAACCGGA  
CAAAATGGAACGAAGTACTCAAAACAAGTCACAGAGAAGCTTAAGGAACATTTCAACAATAAAACCAATTATTTTCAACCCCAAG  
TGCGGAGACCTCGAAATCACTATGACCACTTCAACTGCGCGGGAATTTTATTTGCAATACCACTAAACTTTTCAACAAT  
ACGTGCATCGGAAATGAGACCATGGAGGCTGCAATGGAACAATCATACTCCCATGCAAGATAAAACAATCATTAACATGTGGC  
AAGGTGCTGGACAAGCTATGTATGCACCCCAATATCCGGTAGAATTAATTCGCTCAGCAACATCACTGGCATACTGCTCACTAG  
AGACGGAGGAGCAAAATAATACAAATGAAACATTCGACCCAGGCGGCAACATTAAGGACAACCTGGCGTCCGAACTCTATAAG  
TACAAAGTCGTACAGATCGAACCTCTTGGAATAGCACCGACTCGCGCTAAGACACTCACAGTACAGGCCGACAACTTCTTCTG  
GAATCGTACAGCAATCCAACTCTCTCCGCGCAATCGAGGCCCAACAACATCTGCTTCAGTTCACAGTTGGGGAATCAAGCA  
GCTCCAGGCACCGGTGCTCGCAGTGGAAAGATACCTGAAGATCAGAAATTCCTTGGTCTCTGGGATGTTCTGGCAAAATAATC  
TGCACCTACCGCGGTCCCTGGAATTCACATGGAGCAACCGGAGTTTGAAGAGATATGGAACAATATGACATGGATAGAGTGGG  
AAAGGGAATTAGTAACATACGAACCATGATACGAAATCTCTACCGAAAGCCAAATCAGCAGGATCGCAACGAAAAAGACCT  
CCTCGAGCTTGATAAGTGGGCATCCCTTTTGGAACTGGTTCGACATCACAAATTTGGCTCTGGTaaagatcttataa

57/178

Fig. 34A

Wild-type subtype A Env

00KE\_MSA4076-A (Subtype A, 891 a.a).

MGAMGIQMNWQNLWRWGTMLGMLIICSVAEKSWTVYYGVPVWRDAETTLFCASDAKAHDKVEHNVWATHACVPTDPNPQEMIL  
 ENVTEDFNMWKNMSVEMQHTDIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTKDSATLDMKSEIQNCSEFNMTELRLDK  
 KQKVYSLFYRLDVVQINENSSDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIKP  
 VVTTQLLNGSLAESEVMIRSENITENAKNIIVQFKEPVQIICIRPGNTRKSVHIGPGQAFYATGDIIGDIRQAHNCVSRELWN  
 KTLQEVATQLRKHFERNNTKIIFTNSSGGDEITTHSFNCGGEFFYCDTSGLFNSSWTASNDMSQEAHSTESNITLQCRIKQIINN  
 WQRAGQAMYAPPIPIIRCESNITGLILTRDGGEGNNSTNETFRPVGGNMRDNWRSELYKYKVVEPLGVAPTCSRVRVVEREK  
 RAVGLGAVFIGFLGAAGSTMGAASMTLTVQARQLLSGIVQQSNLLRAIEAQQLHLLKLTWVGIKQLQARVLAVERYLRDQQLLGI  
WGCSGKLICTTNVFPWNSSWSKSLDEIWENMTWMQWDKEVSNTQMIYNLLEESQOQKEQELLALDKWANLWNWENISNWLW  
 YIKIFIMIVGGLIGLRIVFAVLSVINRVROGYSPLSFQHTPNPRGLDRPGRIEEGEGQDRDRSIRLVSGFLALAWDDLRLSLCI  
 FSYHRLRDFILIAARTLELLGHNSLKLRLGWEGLKYLWNLAYWGRELKISALSVDISIAVAGWTDRIIEIVQAGRAILHI  
 PRIRQGLERALI

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 34B

00KE\_MSA4076-A 140CF.pep (647 a.a)

Nick name: 011

MGAMGIQMNWQNLWRWGTMLGMLIICSVAEKSWTVYYGVPVWRDAETTLFCASDAKAHDKVEHNVWATHACVPTDPNPQEMIL  
 ENVTEDFNMWKNMSVEMQHTDIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTKDSATLDMKSEIQNCSEFNMTELRLDK  
 KQKVYSLFYRLDVVQINENSSDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIKP  
 VVTTQLLNGSLAESEVMIRSENITENAKNIIVQFKEPVQIICIRPGNTRKSVHIGPGQAFYATGDIIGDIRQAHNCVSRELWN  
 KTLQEVATQLRKHFERNNTKIIFTNSSGGDEITTHSFNCGGEFFYCDTSGLFNSSWTASNDMSQEAHSTESNITLQCRIKQIINN  
 WQRAGQAMYAPPIPIIRCESNITGLILTRDGGEGNNSTNETFRPVGGNMRDNWRSELYKYKVVEPLGVAPTCSRVRVVEREK  
LLSGIVQQSNLLRAIEAQQLHLLKLTWVGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTNVFPWNSSWSKSLDEIWENMTW  
MQWDKEVSNTQMIYNLLEESQOQKEQELLALDKWANLWNWENISNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

58/178

## Fig. 34C

CODON-OPTIMIZED 00KE\_MSA4076-A 140CF.seq (1972 nt.)

Nick name: 011

ttcagtcgacagccaccatggtgggcaatgggaatccagatgaactggcagaacctcttgccgatggggcacaatgatcctgggtat  
gctcatcatctgctctgttgagaaaagtcattgggtacagctctactacggcgtaaccagtggtggggagcgccgaaacactctc  
ttctggcctccgatgccaaagcacacgataaagaagtcacaaatggttggtgctacccatgctggtgctgcaaccgatcctaacc  
cacaagaaatgatactcgaaaacgttactgaagacttcaacatggtgaaaatcttatggtgaacagatgcacacgacataat  
atcactgtgggatcagctctcaaacctgtgtcaaatgacccccctctggttacactgaactgttccgactcaaatatcact  
tctaattcaacgagcaatagtagcaaacctccgcaaccttgatatgaagaagcaaatacagaactgttcatTTAATATGACCA  
ccgaactgagagataaaagcagaaggtttattctctgttctatcgattggacgtggttcagattAACGAAAATAGCAGCGATT  
ccgactcatTAactgcaatcacacaggttgccccaaaggtAACATTGAGCCAAATCCCTATTCACTACTGCGCC  
cctgcaggatttggccatcctgaaatgcaacgataagaagttTAATGGGACAGGACCCCTGCACCAACGTCTCCACCGTCAATGCA  
cccacggcataaaacctgtgttaccacacaattgctgctcaatggatcaacttgctgaagaggaagtcattgatTCGGTCTGAAA  
catcactgaaaatgccaaaataatatagttcagttcaaaagaacccgtccagatcatTTGCATTGCCCCGTGTAACAACACTCGC  
AAGTCAGTGCAATTGGGCCCGCCAGGCTTTCTATGCAACCGGAGATATATAGGCGACATCAGACAGGCACATTGCAACGTCA  
GCCGGGAATTGTGGAACAAAACCTTTCAGGAAGTTGCTACTCAGCTGCGAAAACATTTAGAAAACAATACAAAGATTATTTTCA  
TAATTCATCAGCGGTGACGTGGAGATCACTACCCATTCAATTAACTGTGGCGGAGAACTCTTATTCGATACCTCTGGGCTC  
TTTAATTCCTCATGGACTGCTAGCAACGATTCAATGCAAGAAGCACATTCACAGAAAGTAATATCACACTGCAGTGCAGAAATTA  
AACAAATCATCAATATGTGCAGCGGCGGTCAGCAATGTAAGCAATGTAAGCACTCCCATCCCCGGAATTAATTCGATGTGAGTCTAATAT  
CACTGGCCTCATTCGACCCGAGACGGTGGCGAAGTAATAATCTACAACAGACTTCAGACCCCTAGGAGGCAATATGCGA  
GACAAATGGCGATCCGAACCTCTGAGCGGAATAGTCCAAACAGCAATCCAATCTTGGAGTGGCACCCCAACCAATCAGGAACCTTGA  
CTGTGAGGCACGCCAATCTGAGCGGAATAGTCCAAACAGCAATCCAATCTTCTGAGAGCTATAGAAGCCAGCAACACCTGCT  
TAACTTACGGTGTGGGAATCAAAACAATTGCAAGCAAGAGTGTGGCAGTGAACGATACCTTGAAGAGCCAAACAATCCTGGGA  
ATCTGGGGATGTTCCGGTAAGTTGATTGCAACGACAAACGTTCCCTGGAACTCTTCTGTCAAAACAAGAGTCTGGACGAAATAT  
GGGAAAATATGACATGGATGCAGTGGGACAGGAAGTTAGCAACTATACAGATGATCTACAACCTCTCGAAGAAATCTCAGAA  
TCAACAGGAAAACGAACAAGAACTGCTCGCCCTCGATAGTGGGCTAACCTCTGGAACTGGTTTAATATTTCAAACTGGTTG  
TGGtaagaagatcttacaa

59/178

Fig. 35A

## Wild-type subtype B

QH0515.1g gp160 (861a.a)

MRVKEIRRNQCRLRRWGTMLLGMLMICSATEQLWVTYYGVPVWKEATTLFCASDAKAYVTEKHNWATHACVPTDPNPQEVVL  
 ENVTFENFMWKNMNMVEQMHEDIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSGTNSSSWEKVQKGEIKNCSEFNITTGIRGRVQ  
 EYSLFYKLDVIPIDSRNNSNNSTEFSSYRLISCNTSVITQACPKEISFEPIPIHYCAPAGFAILKCNCKFNGTGPKCNVSTVQCT  
 HGKIPVSTQLLNGSLAEVEVIRSENFNNVKSIIIVQLNKSVINCTRPNNTRKSIHIGAGKALYTGEIIGDIRQAHCNLSR  
 AQWNNTLKQIVIKLREQFGNKTIVFNQSSGGDVEIVMHSFNCGGEFFYCNSQTQFNSTWNGNDTWKDTTNDNITLPCRIKQ  
 IVNMWQKVGKAMYAPPPIRQIRCSSKITGLILTRDGGTNGTNETETFRPGGNNMKDNWRSELYKYKVVKIEPLGIAPTAKARRVV  
 QREKRAVGTIGAMFLGFLGAAGSTMGAASLTLTQARLLLSGIVQQONNLLRAIEAQOHLQLTVWGIKQLQARVLAVERYLRDQ  
QLLGWGCSGRLICTTNVPWNTSWSNRSNLNIIWDMTMMQWDREINNYTDYIYTLLEDAQNQQEKNEQELLELDKWASLWNWFEDI  
 TNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVQGYSPLSLQTHLPARRGPDPEGIGEGGERDRDRSVRLVHGFLALVWEDL  
 RSLCLFSYHRLRDLILLIVARTVEILQORGWEALKYWNWLLYWSLELKNASVSLVDITIAIAVAEGTDRIEIAARRIFRAFLHIPT  
 RIRQGLERALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design

Fig. 35B

QH0515.1g 140CF (651a.a)

Nick name: 012

MRVKEIRRNQCRLRRWGTMLLGMLMICSATEQLWVTYYGVPVWKEATTLFCASDAKAYVTEKHNWATHACVPTDPNPQEVVL  
 ENVTFENFMWKNMNMVEQMHEDIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSGTNSSSWEKVQKGEIKNCSEFNITTGIRGRVQ  
 EYSLFYKLDVIPIDSRNNSNNSTEFSSYRLISCNTSVITQACPKEISFEPIPIHYCAPAGFAILKCNCKFNGTGPKCNVSTVQCT  
 HGKIPVSTQLLNGSLAEVEVIRSENFNNVKSIIIVQLNKSVINCTRPNNTRKSIHIGAGKALYTGEIIGDIRQAHCNLSR  
 AQWNNTLKQIVIKLREQFGNKTIVFNQSSGGDVEIVMHSFNCGGEFFYCNSQTQFNSTWNGNDTWKDTTNDNITLPCRIKQ  
 IVNMWQKVGKAMYAPPPIRQIRCSSKITGLILTRDGGTNGTNETETFRPGGNNMKDNWRSELYKYKVVKIEPLGIAPTAKAKTLTV  
QARLLSGIVQQONNLLRAIEAQOHLQLTVWGIKQLQARVLAVERYLRDQQLLGWGCSGRLICTTNVPWNTSWSNRSNLNIIWDMTMMQWDREINNYTDYIYTLLEDAQNQQEKNEQELLELDKWASLWNWFEDI  
 TNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVQGYSPLSLQTHLPARRGPDPEGIGEGGERDRDRSVRLVHGFLALVWEDL

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

60/178

## Fig. 35C

CODON-OPTIMIZED QH0515.1g 140CF.seq (1984 nt.)

Nick name:012

ttcagtcgacagccaccatgagagataaaagaaatcagacgcaactgtcagaggttgaggagatgggaaacgatgctcctgggcat  
gctgatgatgttgagtgccaccgaacagctttgggtaacctgtactatggtgtacctgtatgaaagaagccactacaaacctg  
ttttggcggtccgacgcaaaagcctacgtaacagaaaagcacaacgtgtggccacacatgcatgcatgctgccaacagatccaaatc  
ctcaggaagtcgttctggaaaatgtaacagaaaatttttaatatgtggaataaacaataatggtagagcagatgcatgaagatatcat  
ctcactgtgggaacaatcccttgaaaccttggtcaaaccttgaccccaacttgccgtaacacttaactgtactgataagcttcgcaat  
gatacgtccggaaacaattcaagcagctgggaaaagtgcaaaagggcgaaatcaaaaattgttcatttaacatcactaccggta  
tcagagggcggtacaggaataattctcttttcaaaaactcgacgtcatcccaatcgactccagaaataactcaaatataagcac  
agaatttagtagttatcgcccttataagctgcaaacacagcgtgattacacaagcgtgccccataaatctctttgagcccatccct  
attcactactggcaccagccggcttcgccaatcctcaaatgtaacgacaagaaatttaacggaaacggaccctgttaagaaatgtgt  
ccaccgttcaatgcactcatggaatcaagcccgctcgtttctacccaacttcttctcaatggtagccttgccgagaggaaagtgt  
gattcgcctccgaaaattttacaacaacgtcaagtcacatcgtccagcttaataaataccgtcggtatttaattgtacaagaccc  
aacataaacaccagaaaatccattcacatagggcgcggaagctctgtataccggggaattatttgagacatcagacaagcac  
actgtaacttgagtcggcccgagtggaacaacacatggaacagatcgatcaagctcagagagcagttcgggaataagactat  
cgtgtttaatcagagctccgcggtgatgtcgaaatcgtaattgcactcttttaattgtgggggtgaaatttttttactgcaattct  
acacaaattgtttaacagcacctggaacggcaatgacacatggaatgacacctggaaagatacgacaaaatgataataattactcttc  
cgtgcagaataaaagcaaatcgtaaatatgtggcaaaaagtgggcaagccatgtacgcaccacctataagaggacaaattcgctg  
ttcttccaagatcacaggtctgatactcacacgggacggagcgacgaacgggacaaacgagaccgagaccttcggaccaggaggc  
ggcaacatgaaggataactggagaagtgaactttacaagtataaagtgtcaagattgagcctctgggtatcgccccactaagg  
ctaaaacactcacccgtcaggtagattgctgctttcagggatagtcacaacaacagaaacaccttcttagagccattgaagcaca  
acaacacttgcagttgacagtggtgggaattaaacagttgcagggccgggttctcgctgtcgaaacgggtatcttagagatcag  
cagcttttgggtatctgggggtgttcagggccgctcatatgcaccacaatgtcccttggaatacctcatggagtaacaggtctc  
ttaattatatttgggacaatatgacatggatggaatgggataagaaaatttaataactacacccgactacatctacacacttctgga  
ggacggcccaaatcagcaggagaaagacgagcaggaactcctcgaattggataagtgggcatcactgtggaattggttcgatata  
actaatggccttggtaaagatcttaca

61/178

Fig. 36A

Wild-type subtype C

DU123.6 gp160 (854 a.a)

MRVKGIGQNRNWPQWIIWGILGFWMIIICRVVGNLWVTVYGVVPVWTEAKTTLFCASDAKAYEREVHNWVWATHACVPTDPNPQEIIVL  
 GNVTFENFMKNDMVDQMHEDIISWDQSLKPCVKLTPLCVTLNCTDVKNATSNGTITYNNSIDSMNGEIKNCSEFNITTEIRDK  
 KQKVYALFYRPDVVPLNENSSSYILINCENTSTTTQACPKVSFDPIPIHYCAPAGYAILKCNKKTENGTPCHNVSTVQCTHGIRK  
 VVSTQLLNGSLAEELIIRSENLTNNAKTIIVHLNESIEIVCTRPNNTRKSIIRIGPGQTVYATNDIIGDIRQAHCNISKTWN  
 TTLEKVKELKEHFPSKAITTFQPHSGGDLEVTTHSENCERGEFFCYDITKLFNESNLNTTNTTLTPCRIKQIVNMWQGVGRAMY  
 APPVEGNITCNSSITGLLLVRDGGNTSNSTPEIFRPGGMMKDNWRSELYKYKVEIKPLGVAPTAKRRVVEREKRAVGIGAVL  
 EGFAGAAGSTMGAASITLTVOARQLLSGIVQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDDQLLGLWGCSGKLIC  
 PTTVPWNSSWSNKSQTDIWDNMTWMQWDRREISNYTGTIYKLLSESNQOQEKNEKDLLALDSWKNLWSWFDITNWLWYIKIFIMIV  
 GGLIGLRIIFGVLSIVKRVQGYSPLSFTQTLTPNPRGLDRIGRIEEEGEGQDKDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDF  
 ILVAARAVELLGRSSRLGLQRGWEALKYLGNLVQYGGLELKRRAISLEDTIAIAVAEGTDRILEVILRIIRAIRNIPTRIRQGE  
 AALL

Fig. 36B

DU123.6 140CF (638 a.a)

Nick name: 013

MRVKGIGQNRNWPQWIIWGILGFWMIIICRVVGNLWVTVYGVVPVWTEAKTTLFCASDAKAYEREVHNWVWATHACVPTDPNPQEIIVL  
 GNVTFENFMKNDMVDQMHEDIISWDQSLKPCVKLTPLCVTLNCTDVKNATSNGTITYNNSIDSMNGEIKNCSEFNITTEIRDK  
 KQKVYALFYRPDVVPLNENSSSYILINCENTSTTTQACPKVSFDPIPIHYCAPAGYAILKCNKKTENGTPCHNVSTVQCTHGIRK  
 VVSTQLLNGSLAEELIIRSENLTNNAKTIIVHLNESIEIVCTRPNNTRKSIIRIGPGQTVYATNDIIGDIRQAHCNISKTWN  
 TTLEKVKELKEHFPSKAITTFQPHSGGDLEVTTHSENCERGEFFCYDITKLFNESNLNTTNTTLTPCRIKQIVNMWQGVGRAMY  
 APPVEGNITCNSSITGLLLVRDGGNTSNSTPEIFRPGGMMKDNWRSELYKYKVEIKPLGVAPTAKTLTVOARQLLSGIVQQQ  
 SNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDDQLLGLWGCSGKLICPTTVPWNSSWSNKSQTDIWDNMTWMQWDRREISN  
 YTGTYKLLSESNQOQEKNEKDLLALDSWKNLWSWFDITNWLW\*

\*Amino acids seen in blue color is for easy identification of the junction of the

deleted fusion cleavage site.

62/178

## Fig. 36C

CODON-OPTIMIZED DU123.6 140CF.seq (1945 nt.)

Nick name: 013

ttcagtcgacagccaccattgcgcgtaaaggggattcaagaaattggcccgcaattgggtgatttggggaatttctgggcttttgggat  
gataattatatgcccggcttgcggaaatttgcgggtgactgtgtactacggggtgcccgtgtggactgagggcaagaccacccctg  
ttctgtgctagcgatgccaaagccctatgaacgggaagtgcacaattgtttgggctactcatgctgtgtccctaccgaccccaaac  
ctcaggaataatagtcctcgcaatgtaacggaaaacttcaacatgtggaataatgavatggtggtcagatgcacgaagacattat  
ctcaatctgggaccaaagccctgaacccctgacgttaaaactgactcctctcgtcgctcactctcaattgcacagatgtcaaatgaat  
gccacctcaaacggtacgacaaacttaacaacattctattgactctatgaacggcgaaatcaaaaattgttcccttaacatcacca  
ccgagataccggacaaaagcagaaggtctatggccctttttaccgcccgcgacgtagtcctccactcaacgagaattccagctcata  
catcctcaatactgcaatacatcaactaccacacaagcattcccgaaagttagctttgatcccaattcctatacatctactgccc  
ccgcccggctacgctatactgaaatgcaataataagacttttaacgggacccggcccatgtcacacgtgtcaacccgtgcaatgca  
ctcatggcatcaagcccgtggtgcaccagctgctcaatggctcacttgcagaagaagaattattatccgctctgagaa  
tctttactaacaatgcaaaaacgattatcgtgcaccttaatgaatcaatagaatcgtgtgtactcggcccaacaataatactaga  
aaaagcattcgcattcggacctggccagacagtttacgcaactaatgacatcatcggggacatccgacagggccattgcaacattt  
ctaaaaccaagtggaaafacaacccctggaaaagtaagaaaacttaagaaacattttcccttaaggcgatcacgtttcaacc  
tcacagtgccggagacttggaaagtcaaacacattcttttaactgcccgggagaattttttattgtgatacaacaaaacttttt  
aatgaatcaaatcaaacaccacaaatacaaccacactgacctccccctgtagaatcaacaaatcgtaaacatgtggcaagggg  
ttggagggtatgtacgctccccctcgaaaggaataataacgtgtaacagcagcatcactgggctgcttcttgttcgagacgg  
aggcaatacttctaattcaactcctgaaatttttaggcctggcggtggcaataatgaaagataactggcgctcagaactgtacaaa  
tacaaagtgttgaaatttagccctgggagtcgctccaaccaaagctaaaactcacagtgcaagcaagacagctcctttcag  
gcattcgtccagcaacagtc aaatctccttagagcaatcgaaagcccaacagcatatgctcgaactcacagctcggggattaaaca  
gcttcaagccccggcttgctatcgaaacgtatctttaagacccaacagcttcttggcctctggggtgtagtggaataactcattc  
tgccccaccacccgtgcttggaatagttcttggagtaataaatcacagaccgatatattgggacacacatgacctggatgcaatggg  
atagggaatttctaattatactggcacaattctacaaactcttggaaagaaagtc aaatcagcaagaaaaaaacgaaaaggacct  
cctcgccctggactcctggaaagaattcttggagctggttcgacataactaattggctgtggtaaagatcttataa

63/178

Fig. 37A

Wild-type subtype CRF01\_AE

97CNGX2F-AE (854 a.a.)

MRVKETQMNWPNLWKWGTLILGLVLIICSASDNLWVTYYGVPVWRDADTTLFCASDAKAHETEVEHNVWATHACVPTDPNPQEIHL  
 ENVTFENFMRNWNMVEQMVEDVLSLWDQSLKPCVKLTPLCVTLNCTNANWTNSNNTTNGPNKIGNITDEVKNCTFNMTTELKDKK  
 QKVHALFYKLDIVQINSSEYRLINCNTSVIKQACPISFDPIPIHYCTPAGYAILKCNDFNGTGPCKNVSSVQCTHGIKPVS  
 TQLLLNGSLAEIEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITMGPGQVFYRTGDIIGDIRKAYCEINGIKWNEVL  
 VQVTGKLKEHFNKTIIFQPPSGGDLEIITHHFSRCGEFFYCNTTKLFNNTCIGNTSMEGCNNTIILPCKIKQIINMWQGVGQAMY  
 APPISGRINCVSNTIGILLTRDGGADNNTTNETFRPGGNIKDNWRSELYKYKVVEIEPLGIAPTRAKRRVVEREKRAVGIGAMI  
 FGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWCSGKIIC  
 TTAVPWNSSWSNKSFEIWDNMTWIEWEREISNYTSQIYEILTESQNQQDRNEKDLLELDKVASLWNWFDITNWLWYIKIFIIIV  
 GSLIGLRIFAVLSIVNRVRCQYSPLSFQTPTHHQREPDREEIGEAGEQSKDRSVRLVSGFLALAWDDLRLSLCFSYHLLRDF  
 ILIAARTVELLGHSSLKGLRRGWEGLYLGNLLLYWGQEIKISAILLNATAIAVAGWTDREVIEVAQRARALLHIPRRIRQGLE  
 RALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

97CNGX2F-AE 140CF.pep (629 a.a.)

Nick name: 018

MRVKETQMNWPNLWKWGTLILGLVLIICSASDNLWVTYYGVPVWRDADTTLFCASDAKAHETEVEHNVWATHACVPTDPNPQEIHL  
 ENVTFENFMRNWNMVEQMVEDVLSLWDQSLKPCVKLTPLCVTLNCTNANWTNSNNTTNGPNKIGNITDEVKNCTFNMTTELKDKK  
 QKVHALFYKLDIVQINSSEYRLINCNTSVIKQACPISFDPIPIHYCTPAGYAILKCNDFNGTGPCKNVSSVQCTHGIKPVS  
 TQLLLNGSLAEIEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITMGPGQVFYRTGDIIGDIRKAYCEINGIKWNEVL  
 VQVTGKLKEHFNKTIIFQPPSGGDLEIITHHFSRCGEFFYCNTTKLFNNTCIGNTSMEGCNNTIILPCKIKQIINMWQGVGQAMY  
 APPISGRINCVSNTIGILLTRDGGADNNTTNETFRPGGNIKDNWRSELYKYKVVEIEPLGIAPTRAKRRVVEREKRAVGIGAMI  
 SNLLRAIEAQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWCSGKIICCTTAVPWNSSWSNKSFEIWDNMTWIEWEREISN  
 YTSQIYEILTESQNQQDRNEKDLLELDKVASLWNW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

Fig. 37B



64/178

**Fig. 37C****CODON-OPTIMIZED 97CNGX2F-AE 140CF.seq (1921 nt.)****Nick name: 018**

ttcagtcgacàgcccaccatgCGAGTAAAGAGACACAAATGAATTGGCCCAATTTGTGAAGTGGGGAACATTGATCCTGGGACT  
GGTGATAATCTGTAGTGCATCCGACAAATCTCTGGGTGACCGTTTACTATGTTGTACCAAGTTTGGAGAGACGCTGATACCAACCCCTC  
TTCTGTGCAAGCGACGCCAAAGCCACGAAACTGAAGTCCATAATGTATGGGCCACCCACCGTCCGTACCAACCGACCCCTAATC  
CCCAAGAGATCCACCTTGAGAAATGTAACCTGTAACATGTGGAGAAATAACATGGTGAACAATAATGCAGGAAGACGTTAT  
TTCTTGTGGACAGAGCCTTAACCTTGTCAAAATGACTCCCTGTGTGACTCTCAATTTGFACAAACGCAATTTGGACC  
AACAGCAACAACACTACCAACGGCCCTAACAAATTTGGCAATATTACTGATGAAGTCAAGAACTGCACCTTTTAACATGACACAG  
AACTGAAGGATAAGAAACAGAAAGTCCATGCTCTGTCTATAAGCTCGACATAGTACAAATTAATAGCTCAGAAATATAGACTGAT  
AACTGCAATACTTCCGTTATCAAAACAGGCCCTGTCCAAAGATAAGCTTCGATCCCTATTCCTACTGCACACCCAGCCGGT  
TACGCTATCCTGAAATGCAACGATAAGAAATTTAACGGCACAGGTCCTGCARAAACGTTTCTCTGTCCAGTGTACACACGGTA  
TCAAGCCTGTAGTATCAACACAACTGCTCCTGAATGGCTCCTTGGCCGGAAGAGATCATCATTAAGAGTGAGAACCTGACCGAA  
CAACGCCAAGACTATAATAGTGCACCTCAATAAATCTGTAGAAATCAACTGTACCCGACCTCAAAACAACACTCGAACCAAGTATA  
ACAATGGGCCCTGGCCAAAGTTTTTACCGGACCGGACATATAAGCGATATCAGAAAGGCATATTGCGAGATCAATGGCATCA  
AGTGGAAACGAAGTACTGGTTCAAGTAACTGGAAACACTCAAGAACATTTTAATAAGACGATAATATTCAGCCCCCGAGTGGCGG  
CGACCTCGAGATTATCACCCATCACTTTTCTGTAGAGGCGAATTTTTTACTGTAAACACGACCAAGCTCTTCAATAACACGTGC  
ATCGGGAACACTTCTATGGAAGGATGTAATAATACCATTATACTGCCCTGTAAGATCAAGCAGATTATCAACATGTGGCAGGGAG  
TAGGTCAGGCAATGTACGCACCCGATTTTCAGGACGGATCAATTGCGTATCAAAATATCACCGGCATTTCTGTGACCCGGGACGG  
AGCGCAGACAAACATACCACTAACGAGACATTTAGACCTGGAGCGGCAATATAAGGATAATTTGGAGAAGTGAGCTGTATAAA  
TACAAAGTCGTAGATCGAACCCCTCGGCATTGCTCCAACCCGGGCCCGGACTCTACCGTACAAGCTAGACAGCTGCTTTCTG  
GCATAGTCCAACAGCAGTCAAACTCTCCCGCTATTGAAGCACAACAACACCTGCTCCAGCTGACTGTGTGGGAAATCAACACA  
ATTGCAAGCAAGAGTGTGCGCGTGGAAACGCTATTGAAAGATCAGAAATTTCTTGGACTTTGGGGCTGCAGCGGCAAAATATT  
TGTACAACAGCGGTGCTTGGAACTCATCCTGGAGTAATAAAGCTTTGAAGAAATCTGGGACAATATGACATGGATTGAGTGGG  
AGAGAGAGATTTCAAACTATACAAGCCAAATTTACGAAATACTGACAGAAAGTCAAAACCCAGCAGGACAGAAATGAGAAAGACCT  
GCTCGAACTGGATAAGTGGCCCTCTTTGTGGAACTGGGtaagatcttataaa

65/178

Fig. 38A

**Wild-type DRCBL-G (854a.a.)**

MRVKGIQRNWQHLLNWGILILGLVICS~~AEKLWVT~~YYGVVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPOEINMR  
 NVTENFNMWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFNM~~TT~~ELRDKKAEYALFYR  
 TDVVPINEMNNENGTNSTWYRLTNCNVSTIKQACPKVTFEPIPIHYCAPAGFAILKCVDKKFN~~GT~~GTCNNVSTVQCTHGKIPV  
 STQLLNGSLAEKDIIISSENISDN~~AKV~~IIVHLNRSVEINCTRPNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVS~~WT~~KNWNET  
 LRDVQAKLQ~~EY~~FINKSIEFNSSGGDLEITTHSFNCGGEFFYCNTSGLFNNSILKSNISENNDTITLNCKIKQIVRMWQRVGQAM  
 YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGDMKNNRSELYKYTVKIKSLGIAPTRARRVVEREKRAVGVAIF  
 LGFLGTAGSTMGAASITLT~~VQVRQLLSGIVQQSNLLRAIEAQHLLQ~~LTVMGIKQLRARVLALERYLK~~QDQ~~LLGIWGC~~SGK~~LIC  
 TTNPVWNTSWSNKS~~YNEI~~WENMTWIEWEREIDNYTHIYSLIEQSQIQEQKNEQDLLALDQWASLWSFISNWLWYIRIFVMIV  
 GGLIGLRIVFAVL~~SI~~VNRVROGYSPLSFQTL~~LHHQ~~REPDRPAGIEEGGEQDRSIRLVSGFLALAWDDLRLSCLFSYHRLRDF  
 ILIAARTVELLGRNSLKGLRLGWEALKYLN~~LLLY~~WARELKN~~SAINLLDT~~IAIVANWTD~~R~~VIEVAQRA~~AV~~LNIPRRIRQGLE  
 RALL

\*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "w" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "w" will be deleted in 140CF design.

Fig. 38B

**DRCBL-G 140CF.p<sub>ep</sub> (630 a.a.)**

Nick name: 017

MRVKGIQRNWQHLLNWGILILGLVICS~~AEKLWVT~~YYGVVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPOEINMR  
 NVTENFNMWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFNM~~TT~~ELRDKKAEYALFYR  
 TDVVPINEMNNENGTNSTWYRLTNCNVSTIKQACPKVTFEPIPIHYCAPAGFAILKCVDKKFN~~GT~~GTCNNVSTVQCTHGKIPV  
 STQLLNGSLAEKDIIISSENISDN~~AKV~~IIVHLNRSVEINCTRPNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVS~~WT~~KNWNET  
 LRDVQAKLQ~~EY~~FINKSIEFNSSGGDLEITTHSFNCGGEFFYCNTSGLFNNSILKSNISENNDTITLNCKIKQIVRMWQRVGQAM  
 YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGDMKNNRSELYKYTVKIKSLGIAPTRARR~~TLTVQVRQLLSGIVQQQ~~  
 SNLLRAIEAQHLLQ~~LT~~VMGIKQLRARVLALERYLK~~QDQ~~LLGIWGC~~SGK~~LIC~~TTNPVWNTSWSNKS~~YNEI~~WENMTWIEWEREIDN~~  
 YTHIYSLIEQSQIQEQKNEQDLLALDQWASLWSW\*

\*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

66/178

## Fig. 38C

CODON-OPTIMIZED DRCBL-G 140CF.seq (1921 nt.)

Nick name: 017

ttcagtcgacagccaccatgagagttaaaggaaatccaacgcaattggcaacacaccttttggaactggggcataattgattctttggact  
gggtataatttgtagcgctgaaaaaactctgggtactgtctattacggcgctgcctgtctgggagatgccaaacgccccctgttcc  
tgcgcaagtgatgcaaaaggctcacagcactgaatctcaaacatttggggcaccacgcctgtgtgccaacggacccttagtcctc  
aggagatcaacatgagaaacgttacccgaaaatttttaatatgtggaagaaataatatgtggagcaaatgcaggaagacataatttc  
actctgggacgagctctctgaaaccatgtgtgaaacttaccctccctgtgcgtcacctgaaactgtaccgaaatcaacaataactca  
acgagaaatatcacagaagaataaccgaatgactaactgttcccttaatatgacaaccgaaactgcgagacaaaaaagaggctgaat  
acgcacttttctaccgaacagatgttgtaccaatcaacgagatgaacaaatgaaaacaaatggaacgaaactctacctggatagact  
gacaaactgtaacgttagcacaatcaagcagccctgcccctaaagtacacattcgaaaccaataaccaaattcacctactgcccacccgcc  
ggattcgctatttcttaagtgcgtggatagaaagtttaacggaaactggaacctgcaataatgtatctacagtacaaatgcacgcattg  
gaattagcctgtcgtttcaacccagttgctggaatggatcactcgcaaaaaggatatttatctcaagcgaacacatatc  
tgataatgcaaaggctcatcatcgctccacctcaacccgctcagttgaaataaactgcactcgccctaaataataacacagacgctct  
gtcgcaatcgggccaggaacagctttttacactacccgggaagtattcggcgacatacggaaagcccaactgcaacgcttagctgga  
ccaaagtgaatgaaacactgcgcgatgttcaagccaaactcaagaatacttcataaacaatacaattgagttcaattctagctc  
tgcgcgacccctcgagattacaactcactcctttaactgcggcggaattcttttatgtataacctccggtctcttcaacaac  
tctatcctcaaaagtaacatttctgaaaataatgacacaatcacactgaattgcaagatcaagcagattgttaggatgtggcaac  
gagtcggacaaagctatgtacgccccaccatcgccggaaataataacgtgtcgatcaaatatcactggcctcatccttactagaga  
tgccggagacaataatagcaccagcgagataatcagaccagcgcgagatataaaaaacaactggaggtcagagctctacaag  
tacaaaaacagtcaaaattaaaagcctgggcatgtctccactcgggcccgccacactgactgtccaagtccgacagctcctgtccg  
gaatcgctccaaacagtcacacttgcctggcgctatagaggctcaacaacatctccttcaactgactgtgtggggatcaaaaca  
attgagagcaagagtgctggcgctggaacggatctttaaggaccacaactcctgggcataatgggggtgtccggcaaacctgac  
tgcacaacaaatgtaccttggaacacacagctggtcaataaaaattataatgagatgggaaacacatgacatggattgaatggg  
aaagggaattgacaaattatacataccatataactctctcatcgaaacaaatctcagatacaacacaggaataagaatgaacaagattt  
gttggtctcttgacccaatgggcttctttgtggagtttgtaagatcttaca

## 2003 Centralized HIV-1 Envelope Proteins and the Codon-Optimized Gene sequences

Fig. 39A

## 2003 Cons Env

MRVMGIQRNCQHLWRWGILIFGMLIICSAEENLWTVVYGVVPVWKEANTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIIVLENTENF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTNNEEIKNCSEFNITTEIRDKKKVYALFYKLDVVPIDNNNSYRLI  
 NCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEIIIRSENITNNAKTIIV  
 QLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISRTKWNKTLQQVAKKLREHFNKTIIFNPSSGGDLEITTHSFNCGGE  
 FFYCNSTSELFNSTWNGTNNITITLPCRIKQIINMWQGVQAMYPPIEGKIRCTSNITGLLLTRDGGNNNTETFRPGGGDMRDNRSELYKYK  
 VVKIEPLGVAPTAKRRVVEREKRAVGIGAVFLGAGSTMGAASITLTQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQAR  
 VLAVERYLKDQQLGIWCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMEWDKEINNYTDIIYSLIEESQNOQEKNEQELLALDKWASLWN  
 WFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDPRPEGIEEGEGQDRDRSIRLVNGFLALAWDDLRSL  
 CLFSYHRLRLDLILIAARTVELLGRRGWEALKYLWNLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRCRAILNIPRRIRQGFERAI  
 LLS

Fig. 40A

## 2003 M. Group. Anc. Env

MRVMGIQRNCQHLWRWGILIFGMLIICSAEENLWTVVYGVVPVWKEANTTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIIVLENTENI  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTNNGEIKNCSEFNITTEIRDKKKVYALFYRLDVVPIDNNNSYRLI  
 NCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEIIIRSENITDNAKTIIV  
 QLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGAENKTLQQVAAKLREHFNNKTIIFKPSSGGDLEITTHSFNCGG  
 EFFYCNSTGLFNSTWNGTNETITLPCRIKQIVNMWQVRVQAMYPPIAGNITCKSNITGLLLTRDGGTNTTETFRPGGGDMRDNRSELYKY  
 KVVIEPLGVAPTAKRRVVEREKRAVGIGAVFLGAGSTMGAASITLTQARQLLSGIVQQSNLLRAIEAQHLLQLTVWGIKQLQA  
 RVLAVERYLKDQQLGIWCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMEWDKEINNYTDIIYSLIEESQNOQEKNEQELLALDKWASLW  
 NWFDTNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDPRPGIEEGEGQDRDRSIRLVSGFLALAWDDLRSL  
 LCLFSYHRLRLDLILIAARTVELLGRRGWEALKYLWNLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRCRAILNIPRRIRQGFERA  
 LLS

67/178

69/178

Fig. 40B

## 2003 M. Group. anc Env. seq. opt

ATGCGGTGATGGCATCCAGCGCAACTGCCAGCACCTGTGGCGTGGGGCATCTTGATCTTCGGCATGCTGATGATCTGCTCCGCCGCCGA  
GAACCTGTGGTGACCGTGAACGCGTGGCGTGGAGAGGCCAACACCCACCTGTTCTGGCTCCGACGCCAAGGCCCTACGACA  
CCGAGTGCACAACTGTGGGCCACCCACGCTCGTGGCCACCGACCCCAACCCAGGAGATGCTGTGGAGAACCTTC  
AATGTGGAAGAACACATGTTGGAGCAGATGCAGGAGACATCATCTCCCTGTGGGACCATCTCTGAAGCCCTGCGTGAAGCTGACCCC  
CCTGTGCGTGACCTGACTGACACCGAGTGAACGCCACCAACAACTCCACCAACATGGGCGAGATCAAGAACTGCTCTCAACATCAGCA  
CCGAGATCCGCGACAAAGACAGAGGTGTACGCCCTGTTCTACCGCTGGAGTGGTCCCATCAAGACAACAACTCTACGCCCTGATC  
AACTGCAACACCTCCGCCATCACCCAGGCCCTGCCCAAGGTCTCTTCAGGCCATCCCATCTACTTGGCCCCCGGCGGCTTCGCCAT  
CCTGAAGTGAACGACAGAAATTCAAGGCAACGGCCCTGCAAGAACGTGTCCACCTGCACTGAGTGCAGTGCACCCACGGCATCAAGCCGTGGTGT  
CCACCCAGCTGCTGTGACCGCTCCCTGGCCGAGGAGATCATCATCCGCTCCGAGAACATCAACGACAACGCCAAGACCATCATCGTG  
CAGCTGAACGAGTCCGTGGAGATCAACTGCACCGGCCCAACAAACACCGCAAGTCCATCCGATCGGCCCGGCCAGGCCCTTCTACGC  
CACCGCGACATCATCGCGACATCCGCCAGGCCACCTGCAACATCTCCGGCGCGAGTGGAAACAGACCTGCAGAGGTGGCCGCCAAGC  
TGCGGAGCACTTCAACAAAGACCATCATCTCAAGCCCTCTCCGGCGCGGACCTGGAGATCAACCCACTCTTCAACTGCGGCGGC  
GAGTCTTCTACTGCAACACCTCCGGCTGTTCACTCCACCTGGAACGGCACCAACGAGACCATCACTGCCCTGCCGATCAAGCAGAT  
CGTGAACATGTGGCAGCGGTGGCCAGGCCATGTAGCCCCCCCCATCGCCGSCAACATCACCTGCAAGTCCAACATCACCGGCCCTGCTG  
TGACCCGCGACGCGGCACCAACAAACACCGAGACCTTCGCCGCCGCGGCGACATGCGCGACAACCTGGCGCTCCGAGCTGTACAAGTAC  
AAGGTGTTGAAGATCGAGCCCTTGGCGTGGCCGCCCAACAGGCCAAGCGCGGTGGTGGAGCGGAGAACGCCGCCGTGGGCATCGGCG  
CGTGTCTCTGGCTTCTTGGCGCGCGCGGCTCCACCTAGGCGCGCGCTCCATCACCTGACCTGACCTGACCGTGGCGCGCGAGCTGTCCGGC  
TCGTGCAGCAGCATCCAACTGCTGCGCGCCATCGAGGCCAGCAGCACCTGCTGACGTGACCTGCTGGGCGATCAAGCAGCTGCAGGC  
CGCGTGTGCGCGTGGAGCGCTACCTGAAGGACACGAGCTGCTGGGCTGCTCCGGCAAGTGTATGACCAACCAACGTGAC  
CTGGAACCTCTCTGTCCAAAGTCCAGGACGAGATCTGGGACAACATGACCTGGATGCACTGGGAGCGGAGATCTCAACTACACCG  
ACATCATCTACTCCCTGATCGAGGAGTCCAGAACACGAGGAGAGAACGAGACGAGACCTGCTGGCCCTGGACAAGTGGGCCCTCCCTGTGG  
AACTGGTTCGACATACCAACTGGCTGTGTACATCAAGATCTTATCATGATCGTGGCGGCGCTGATCGGCCCTGGCATCGTGTTCGCCGT  
GCTGTCCATCGTGAACCGCTGCGCCAGGCTACTCCCCCTGTCTTCCAGACCTGTATCCCAACCCCCCGGCCCGGACCGCCCGGGG  
GCATCGAGGAGGCGCGGAGCAGGACCGGACCGCTCCATCCGCTGGTCTCGGCTTCTTGGCCCTGGCCCTGGGACGACCTGCGCTCC  
CTGTGCCCTGTTCTCTACCCAGCCCTGCGCGACTTCACTCTGATCGCCGCCGACCGTGGAGCTGTGGGCCCGCGCGGCTGGGAGGCCCT  
GAAGTACCTGTGGAACCTGTGCACTGAGTACTGGGCGCAGGAGTGAAGAACTCCGCGCATCTCCCTGCTGGACACCAACCGCATCGCCGTGGCCG  
AGGGCACCGCGGTGATCGAGGTGGTGCAGCGCGCTGCCCGGCCCATCTCTGCACATCCCCCGCGCCGATCCGCCAGGGCTTCGAGCGCGCC  
CTGCTGTAA

70/178

Fig. 41A

2003 CON A1 Env

MRVMGIQRNCQHLLRWGTMILGMIIICSAENLWVTYYGVVWKDAETTLFCASDAKAYETEMHNWATHACVPTDPNPQEIHLNVTEEF  
 NMWKNMVEQMHADIISLWDQSLKPCVKLTPLCVTLNCSNVNVTNNTNTHHEEIKNCSENMTELDRDKKQVYSLFYRLDVVQINENNSNS  
 SYRLINCNVTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEVIRSENITNNA  
 KTIIVQLTKPVKINCTRPNNNTRKSIIRIGPGQAFYATGDIIGDIRQAHNCNVRSEWNKTQKVAKQLRKYFKNKTIIIFTNSSGGDLITTHS  
 FNCGGEFFYCNVTSGLFNSTWNGTMKNTITLPCRKQIINMWQAGQAMYPPIQGVIRCESNITGLLTRDGGNNNTNETFRPGGDMRDN  
 WRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLKLT  
 WGIKQLOARVLAVERYLKDQQLGIWGCCKLICTTNVFNWSSWSNKSQNEIWDNMTWLDWKEISNYTHIIYNLIEESQOQEKNEQDLLA  
 LDKWANLWNWFDISNWLWYIKIFIMIVGGLIGLRIVEAVLSVINRVQGYSPLSFQHTPNPRGLDRPGRIEEGEGQGRDRSIRLVSGFLA  
 LAWDDLRLSLCLFSYHRLRDFILIAARTVELLGHSSLKGLRLGWEGLYLWNLNLLYWGRELKISAINLVDITIAIAGVWTDRIEIGRICRA  
 ILHIPRRIRQGLERALL\$

Fig. 42A

2003 A1.Anc Env

MRVMGIQRNCQHLLRWGTMILGMIIICSAENLWVTYYGVVWKDAETTLFCASDAKAYDTEVHNWATHACVPTDPNPQEIHLNVTEEF  
 NMWKNMVEQMHADIISLWDQSLKPCVKLTPLCVTLNCSNVNVTNNTNTHHEEIKNCSENMTELDRDKKQVYSLFYRLDVVQINENNSNS  
 SYRLINCNVTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEVIRSENITNNA  
 KTIIVQLTKPVKINCTRPNNNTRKSIIRIGPGQAFYATGDIIGDIRQAHNCNVRSEWNKTQKVAKQLRKHFNKNTIIFNSSGGDLITTHS  
 FNCGGEFFYCNVTSGLFNSTWNGTMKNTITLPCRKQIINMWQVQAMYPPIQGVIRCESNITGLLTRDGGNNNTNETFRPGGDMRDN  
 WRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQHLLKLT  
 WGIKQLOARVLAVERYLKDQQLGIWGCCKLICTTNVFNWSSWSNKSQDEIWDNMTWLDWKEISNYTHIIYNLIEESQOQEKNEQDLLA  
 LDKWANLWNWFDISNWLWYIKIFIMIVGGLIGLRIVEAVLSVINRVQGYSPLSFQHTPNPRGLDRPGRIEEGEGQGRDRSIRLVSGFLA  
 LAWDDLRLSLCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGLYLWNLNLLYWGRELKISAINLDTIAIAGVWTDRIEIGRICRA  
 ILNIPRRIRQGLERALL\$



[illegible]



68/178

Fig. 39B

## 2003 CON-S Env. seq. opt

ATCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCATCTGTGATCTTTCGGCATGCTGATCATCTGCTCCGCCGCCCGA  
GAACCTGTGGTGACCGTGTAACCGCGTGCCCGTGTGGAAGAGGCCAACACACACCTCTGTTCGGCGCTCCGACGCCAAGGCCCTACGACA  
CCGAGGTGCACAACTGTGGGCCACCCACCGCTGCTGCCACCCGACCCCAACCCAGAGAGATCGTGTGGAGAACGTGACCGAGAACTTC  
AACAATGGGAAGAACAACTATGTTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCACTCCCTGAAGCCCTGCGTGAAGTGAACCC  
CCTGTGCGTGACCCCTGAACCTGCACCGGACGTAACCGCCACCAACACCAACAGGAGATCAAGAACTGCTCTTCAACATCAACCA  
CCGAGATCCGCGACCAAGAAAGGTGTACGCCCCTGTCTACAAGCTGGACGTGGTCCCATCGACGACAACAACTCCTACCGCCTGATC  
AACTGCAACACCTCCGCCATCACCCAGGCTGCCCAAGGTGTCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGCCAT  
CCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCTTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGT  
CCACCCAGCTGTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCTCCGCTCCGAGAACATCACCAACAACGCCCAAGACCATCATCGTG  
CAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCAACAACACCCGCAAGTCCATCCGCATCGGCCCGCGCAGGCCCTTCTACGC  
CACCGCGACATCATCGCGGACATCCGCCAGGCCCACTGCAACATCTCCCGCACCAAGTGAACAAGACCCCTGCAGCAGGTGGCCAAAGAAGC  
TGCGGAGCACCTTCAACAAGACCATCATCTTCAACCCCTCTCCGGCGGACCTGGAGATCACCAACCTCTTCAACTGCGCGCGGAG  
TTCCTTCTACTGCAACACCTCCGAGCTGTCAACTCCACCTGGAACGGCACCAACAACACCATCACTCCCTGCCCTGCCGATCAAGCAGATCAI  
CAACATGTGGCAGGGCTGGSCCAGGCCATGTACGCCGCCCTCCATCGAGGGCAAGATCCGCTGCACCTCCAAACATCACCGGCCCTGCTGTG  
CCCGCACGGCGGCAACAACAACACCGAGACCTTCCGCCCGCGGGCGGACATGCGCGACAACTGGCGCTCCGAGCTGTAACAAGTACAA  
GTGTTGAAGATCGAGCCCTGGCGGTGGCCCCCAACAGGCCAAGCGCCGCTGTGGTGAAGCGGAGAACCGGCCCTGGGATCGGGCCCG  
GTTCTGGGCTTCTTGGGCGCCCGGCTCCACCATGGGCGCCGCTCCATCACCTGACCGTGCAGGCCCGCCAGCTGTCCGGCATCC  
TGACGACGAGTCCAACTGTGCGCGCCATCGAGGCCCGCAGCACCTGTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGC  
GTGCTGGCGGTGGAGCGCTACCTGAAGGACCAAGCTGTGGGCATCTGGGCTGTCCGGCAAGCTGATCTGCACCAACCAAGTGCCTG  
GAACCTCTCTGGTCCAAACAGTCCAGGACGAGATCTGGGACAAACATGACCTGGATGGGACAAAGGAGATCAACAACATCAACCGACA  
TCATCTACTCCCTGATCGAGGAGTCCAGAACCGACGAGAGAACGAGAGTGTGGCCCTGGACAAGTGGCCCTCCCTGTGGAAC  
TGGTTCGACATCAACCAACTGGCTGTGGTACATCAAGATCTTCAATGATCGTGGCGGCTGTATCGGCTGCGCATCGTGTTCGCCGTGCT  
GTCCATCGTGAACCGGTGCGCCAGGCTACTCCCGCTGTCTTCCAGACCTGTATCCCAACCCCGCGGCCCGACCGCCCGAGGGCA  
TCGAGGAGGAGGGCGGAGACCGGACCGCTCCATCCGCTGTGGTGAACGGCTTCTGGCCCTGGCCCTGGGACGACCTGGCTCCCTG  
TGCCCTGTTCTCTACACCGCTGCGGACCTGATCCTGATCGCCCGCCGACCGTGGAGTGTGGCCCGCGCGGCTGGAGGCCCTGAA  
GTACCTGTGGAACCTGTGAGTACTGGGGCCAGGAGTGAAGAACTCCGCCATCTCCCTGTGGACACCAACCGCCATCGCCGTGGCCGAGG  
GCACCGACCGGTGATCGAGGTGGTGCAGCGCGTGTGCCCGGCCATCTCTGAACATCCCCCGCCGATCCGCCAGGGCTTCGAGCGCGCCCTG  
CTGTAA

73/178

Fig. 43A

2003 CON A2 Env

MRVMGTQRNYQHLLWRWGILILGMLIMCKATDLWVTYYYGVVPWKDADTTLCASDAKAYDTEVHNWVWATHACVPTDPNPQEVNLENVTEDFN  
 MWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCSNANTNNSTMEIKNCSYNITTELKDKTKQVYSLEYKLDVVQLDESNKSEYYR  
 LINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDPRFNGTGCNNVSSVQCTHGKIPVASTQLLNGSLAEGKVMIRSENITNNAKNI  
 IVQFNKPVPITCIRPNNTKRSIRFGPGQAFYTNDIIGDIRQAHCNINKTKWNATLQKVAEQLEHFPNKTIIFTNSSGGDLEITTHSFNCG  
 GEFFYCNTTGLFNSWKNGTNNTEOMITLPCRKQIINMWQRVGRAMYAPPIAGVIKCTSNITGIILTRDGGNNETETFRPGGGDMRDNR  
 SELYKVKVVKIEPLGVAPTRAKRRVVEREKRAVGMGAVFLGFLGAAGSTMGAASITLTVOARQLLSGIVQQQSNLLKAIEAQHLLKLTVMG  
 IKQOARVLALERYLDQQLLGIWGCCKLICATTVPWNSSWSNKTQEEIWNNTWLQWDKEISNYTNIYKLLSESONQQEKNEQDALLD  
 KWANLWNWENITNWLWYIRIFIMVGGGLIGLRIVIAIISVNRVRQGYSPLSFQIPTNPEGLDRPGRIEEGGEGQGRDRSIRLVSGFLALA  
 WDDLRLSLCLFSYHRLRDCILIAARTVELLGHSSLKGLRLGWEGLKYLWNLLLYWGRELKNSAISLLDTIAVAVAEWTDRIEIGQACRAIL  
 NIPRRIRQGFERALL\$

Fig. 44A

2003 CON B Env

MRVKGIRKNYQHLLWRWGTMILGMLIMCSAAEKLWVTYYYGVVPWKAEATTLFCASDAKAYDTEVHNWVWATHACVPTDPNPQEVNLENVTENF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDEMNAATNTTIIYRWGEIKNCSEFNITTSIRDKVQKEYALFYKLDVVPIDND  
 NTSYRLISCNSTSVITQACPKVSFEPIPIHYCAPAGFAILKCKDPRFNGTGCNNVSSVQCTHGKIPVASTQLLNGSLAEGKVMIRSENFTD  
 NAKTIIIVQLNESVEINCTRPNNNTKRSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLQIVKKLREQFGNKTIVFNQSSGGDPEIVM  
 HSFNCGGEFFYCNTTQLFNSTWNGTWNNTGNITLPCRKQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLLLTRDGGNNETETFRPGGGDM  
 RDNWRSELYKYKVKVVKIEPLGVAPTRAKRRVVEREKRAVGMGAVFLGFLGAAGSTMGAASMTLTVOARQLLSGIVQQQSNLLKAIEAQHLLQ  
 LTVWGIKQOARVLAVERYLDQQLLGIWGCCKLICATTVPWNASWSNKSLEIWDNMTWMEWEREIDNYTSLIYTLIEESQNEQKNEQE  
 LLELDKWASLWNWFDITNWLWYIKIFIMVGGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTRLPAPRGPDRPEGIEEGGERDRDRSGRLVDG  
 FLALIWDCLRSLCLFSYHRLRDLILLIVTRIVELLGRGWEVLKYWNNLLQYWSQELKN\$AVSLNATAIAVAEGTDRVIEVQVQACRAILLHI  
 PRRIRQGLERALL\$

74/178

## Fig. 43B

2003 CON A2 Env. seq. opt

ATCGCGGTGATGGGCAACCCAGCGCAACTACAGCACCTGTGGCGTGGGGCATCTTGATCCTGGGCATGCTGATCATGTGCAAGGCCACCCGA  
CCTGTGGGTGACCGTGTACTACGGCGTCCCGTGTGGAAAGAGCGCGGACACACCTCTGTTCTGGCCCTCCGACGCCAAGGCCCTACGACACCG  
AGGTGCACAACGTGTGGGCCACCCACCGCTGCGTCCACCGACCCCAACCCAGGAGGTGAACCTGGAGAACGTGACCGAGACTTCAAC  
ATGTGAAGAACAACATGCTGGAGCAGATGCACGAGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCCTTCAAC  
GTGCGTGACCCCTGAACCTGCTCCAAAGCCCAACACCAACCAACTCCACCATGGAGGAGATCAAGAACTGCTCCCTACAACATCACCAACCGAGC  
TGCGGACAAAGACCCAGAGGTGTAATCCCTGTTCTACAAGCTGGACCTGGTGCAGTGGACGAGTCCAACAAGTCCGAGTACTACTACCGC  
CTGATCAACTGCAACACCTCCGCCATCACCGAGCTGCCCAAGGTGTCTCTCGAGCCCATCCCATCCACTACTGCGCCCTCCGCGGCTT  
CGCCATCCTGAAGTGAAGACCCCGCTTCAAGGCAACGGCTCTGCAACAACAGTGTCTCCGTGCACTGCAACCAACGCGCATCAAGCCG  
TGGCCTCCACCCAGCTGCTGTGAACGGCTCCCTGGCCGAGGGCAAGGTGATCCGCTCCGAGAACATCAACAACAACGCGCATCAAGCAATC  
ATCGTGCAGTTCAACAAGCCCGTGCCTCATACCTGCATCCGCCCAACAACAACAACCCGCAAGTCCATCCGTTCCGCCCTCCGCGCAGGCCTT  
CTACACCAAGACATCATCGCGGACATCCGCCAGGCCCACTGCAACATCAACAAGACCAAGTGGAAACGCCACCTCGAGAAAGTGGCCGAGC  
AGCTGGCGGAGCACTTCCCAACAAGACCATCATCTTCACTCACTCCCGCGGCGACCTGGAGATCACCAACCTCCCTCAACTGCGGC  
GGCGAGTTCTTACTGCAACACCAACCGGCTGTTCACCTCGAAGAACGGCACCAACCAACAACCGAGCAGATGATCACCCCTGCC  
CTGCCGATCAAGCAGATCATCAACATGTGGCAGCGGTGGCCCGCCATGTACGCCCGCCCATCGCCCGCGTGTCAAGTGCACTCCA  
ACATCACCGGCATCATCTGACCCCGGACGGGCAACAACGAGACCGAGACCTTCCGCCCGCGCGGCGGACATGCGCGACAACTGGCGC  
TCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTGGCGGTGGCCCGCCACCGCGCAAGCGCGCTGGTGGAGCGGAGAAGCG  
CGCGTGGCATGGCGCGCTGTCTGGCTTCCCTGGCGCGCGCGCTCCACCATGGCGCGCGCTCCATCACCTGACCGTGCAGGCCC  
GCCAGTGTGTCCGGCATCGTGCAGCAGTCCCAACCTGCTGAAGCCATCGAGGCCCGAGCACCTGCTGAAGCTGACCGTGTGGGC  
ATCAAGCAGTGCAGGCCCGCTGTCCTGAGCGCTACCTGCAGGACCGAGCTGTGGGCATCTGGGCTGCTCCGCGCAAGCTGAT  
CTGGCCACCCGCTGGAACTCCTCTGTTCCAAAGACCCAGAGGAGATCTGGAACAACATGACCTGGCTGCAGTGGGACAAGG  
AGATCTCCAACTACCAACATCATCTACAAGCTGTGGAGGAGTCCCAAGAACGAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGAC  
AAGTGGCCAACTGTGGAACTGGTTCAACATCAACCACTGGCTGTGGTACATCCGCACTTCACTCATGATCGTGGCGCGCTGATCGGCCT  
GGCATCGTGATCGCCATCATCTCCGTGGTGAACCGCGTGGCGCAGGGCTACTCCCCCTGTCTTCCAGATCCCCAACCCCAACCCGAGG  
GCCGTGACCCCGCGCATCGAGAGGGCGGGCGGAGAGGGCCCGACCGCTCCATCCGCTGGTGTCCGGCTTCTGGCCCTGGCC  
TGGGACGACCTGCGCTCCCTGTGCCTGTTCTCTACCAACCGCTGCGCGACTGCATCTGATCGCGCCCGCACCGTGGAGCTGCTGGGCCA  
CTGCTCCCTGAAGGGCTGCGCCTGGGCTGGAGGGCTGAAGTACCTGTGGAACCTGCTGTACTGGGCGCGGAGCTGAAGAACTCCG  
CCATCTCCCTGCTGGACACCATCGCCGTGGCGGTGGCCGAGTGGACCGCGCTGATCGAGATCGGCCAGCGCGCTGCGCGCCATCCTG  
AACATCCCCCGCGCATCCGCCAGGGCTTCGAGCGCGCTGCTGTAA

75/178

Fig. 44B

2003 CON B Env. seq. opt

ATGCGGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTGGCGTGGGGACCATGCTGCTGGGCATGCTGATGATCTGCTCGGCCGCCGA  
 GAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGAAGAGGCCACCAACCCCTGTTCTGCGCTCCGACGCCAAGCCCTACGACA  
 CCGAGGTGCACAACGTTGGGCCACCCACCGCTGCGTCCCAACCCCAACCCAGGAGTGTGCTGGAGAACGTGACCGAGAACTTC  
 AACATGTGAAGAACAACTGTGGAGCAGATGCACGAGACATCATCTCCCTGTGGGACAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCCCTGAATGCACCGACCTGATGAACGCCACCAACACCAACCATCATCTACCGCTGGCGGGGAGATCAAGAAT  
 GCTCCTTCAACATCACCACTCCATCCGGACAAGTGCAGAGGATAGCCCTGTTCTACAAGCTGGACGTGGTGGCCATCGACAACGAC  
 AACACCTCTACCGCTGATCTCCTGCAACACCTCCGTGATCACCCAGGCTGCCCAAGGTCTCTCGAGCCCATCCCATCCACTACTG  
 CCCCCCGCGGCTTCGCCATCCTGAAGTGCAACGACAAAGATTCAACGGCACCGGCCCTGCACCAACGTGTCCACCGTGCAGTGCACCC  
 ACGGCATCCGCCCCGTGGTGTCCACCCAGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGGTGGTGTCTCGCTCCGAGAACTTCAACGAC  
 AACGCCAAGACCATCATCGTGCAAGTGAACGAGTCCGTGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACATCGG  
 CCCCCCGCGGCTTCTACACCAACCGCGAGATCATCGCGACATCCGCGACATCCGCGAGGCCCACTGCAACATCTCCGCGCCAAAGTGAACAACACCC  
 TGAAGCAGATCGTGAAGAGCTGCGGAGCAGTTCCGCAACAAGACCATCGTGTCAACCACTCCCGCGGCGACCCCGAGATCGTGATG  
 CACTCCTTCAACTGCGCGGCGAGTTCTTACTGCAACACCAACCCAGCTGTTCAACCACTCCCTGGAACGGCACCTGGAACAACACCCGAGT  
 CAACATCACCTGCCCTGCCGATCAAGCAGATCATCAAGTGTGAAGTGTGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACATCGG  
 TCCGCTGCTCCTCCAACATCACCGAGCTGTACAAAGTACAAAGTGTGAGATCGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACATCGG  
 CGCACAACCTGGCGCTCCGAGCTGTACAAAGTACAAAGTGTGAGATCGAGATCAACTGCACCCGCCCAACAACAACACCCGCAAGTCCATCCACATCGG  
 GCAGCGGAGAGCGCGCGTGGCATCGCGATCGCGCATCGTGGCTTCTGCGCTTCTGGCGCGCGCGCTCCACCATGGCGCGCGCTCCATGACCC  
 TGACCGTGACGCGCGCGCATCGTGCAGCAGAACACCTGCTGCGCGCATCGAGGCCACAGCAGTGTGGGCAACATCGTGGGCACTTGGGGCTG  
 CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCGCGCGTGTGGCGTGAGCGCTACCTGAAGGACCAAGCAGTGTGGGCACTTGGGCAACATGACCTGGA  
 CTCCGGCAAGCTGATCTGCACCAACCGCGTGCCTGGAAACGCTCCTGGTCCAACAAGTCCCTGGACGAGATCTGGGCAACATGACCTGGA  
 TGGAGTGGGAGCGGAGATCGAACAACCTCCCTGATCTACACCTGATCGAGGAGTCCCAAGAACCAAGCAGGAGAGAACGAGCAGGAG  
 CTGCTGGAGCTGACAAGTGGCTTCCCTGTGGAACTGGTTCGACATCAACAAGTGGTGTGATCAAGATCTTCAATCATGATCGTGGG  
 CGGCTGGTGGCTTGGCATCGTGTGCGCGTGTGTCATCGTGAACCGCGTGGCGAGGCTACTCCCGCTGTCTTCCAGACCCGCGC  
 TGCCCGCCCCCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGCGGCGAGCCGCAACCGTCCGCGCGCTGGTGGACGGC  
 TTCTGGCCCTGATCTGGGACGACCTGGCGTCCCTGTGCTGTTCTTCTACCAACCGCTGCGCGACCTGTGCTGATCGTGACCCGATCGT  
 GGAGTGTGGGCGCGCGGCTGGAGGTGCTGAAGTACTGTGGAACCTGTGCTGAGTACTGTTCCAGGAGCTGAAGAATCCGCGCTGT  
 CCCTGCTGAACGCCACCGCCATCGCCGTGGCGGAGGCAACCGCGTGTGATCGAGGTGGTGCAGCGCGCTGCCGCGCATCTCTGCACATC  
 CCCCCCGCATCCGCGCAGGGCTTGAAGCGCGCGCTGCTGTAA

76/178

Fig. 45A

2003 B.anc Env

MRVKGIRKNCQHWRWGTMLLGMMLICSAENLWTVVYGVVPWKEATTLLFCASDAKAYETEVEHNVWATHACVPTDPNPQEVVLENVTEF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDNLNATNTSTNMYRWRGEIKNCSENIITTSIRDKMQKEYALFYKLDVVPIDNN  
 TSYRLINCNTSVITQACPVSFEPPIPIHYCTPAGFAILKCNCKFNGTGPCKNVSTVQCTHGIRPVVSTQLLNGSLAEEVVIRSENFDTN  
 AKTIIVQLNESVEINCTRPNNTRKSIHIGPGRAFYATGEIIGDIRQAHCNLSRAKWNNTLKQVVTKLREQFDNKTIVFNPSSGGDPEIVMH  
 SFNCGGEFFYCNTTQLENSTWNGTWNTEGNTITLPCRKIQIINMWQEVGKAMYAPPPIRGQIRCSSNITGLLTRDGGNNETEIFRPGGDMR  
 DNWRSELYKYKVVKIEPLGVAFTKAKRRVVQREKRAVGIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQNNLLRAIEAQHLLQL  
 TVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTTPWNASWSNKSLSDEIWNMTWMEWEREIDNYTGLIYTLIEESQOQEKNEQEL  
 LELDKWASLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTRLPA PRGPD RPGEIEEGGERDRDRSGRLVNGF  
 LALIWDRLSLCLFSYHRLRDLILLIVARIVELLGRRGWEALKYWNLLQYWSQELKNSAVSLINATAIAVAEGTDRVIEVVQACRAILHIP  
 RRIRQGLERALL\$

Fig. 46A

2003 CON C Env

MRVRGILRNCQQWIIWGIWGFWMLMNCVNVGNLWTVVYGVVPWKEAKTLLFCASDAKAYEKEVHNVWATHACVPTDPNPQEVVLENVTEF  
 NMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATMTMGEIKNCSENIITTEL RDKKQKVYALFYRLDIVPLNENNSYRLINC  
 NTSAITQACPVSFDPPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGIRPVVSTQLLNGSLAEEIIIRSENLTNNAKTIIVHL  
 NESVEIVCTRPNNTRKSIHIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLOKVSKKLKEHFPNKTIKFEPSSGGDLEITTHSFNCRGEF  
 FYCNTSKLENSTYNSTNTITLPCRKIQIINMWQEVGRAMYAPPIAGNITCKSNITGLLTRDGGKNNNETFRPGGGDMRDNWRSELYKYKV  
 VEIKPLGIAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTQARQLLSGIVQQSNLLRAIEAQHMLQLTVWGIKQLQTRV  
 LAIERYLKQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEDIDWNMTWQWDREISNYTDTIYRLLEDSONQOQEKNEKDLALDLSWKNLWNW  
 FDI TNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPNPRGPDRLGRIEEGGEQDRDRSIRLVSGFLALAWDDLRLSLC  
 LFSYHRLRDFILIAARAVELLGRSSRLRGLQRGWEALKYGLSLVQYWGLELKSALSLDITIAIAVAEGTDRIIELIQICRAIRNIPRRIRQ  
 GFEAALQ\$

77/178

## Fig. 45B

## 2003 B.anc Env. seq. opt

ATGCGGTGAAGGCGCATCCGCAAGAACTGCCAGACCTGTGGCGCTGGGGCACCACATGCTGTGGGCATGCTGATGATCTGCTCCGCGCGCCGA  
 GAACCTGTGGTGACCGTGTAACAGGCGTGCGGTGGAAGGAGCCACACACACCTGTTCTGGGCTCCGACGCCAAGGCCTACGAGA  
 CCGAGGTGCACAACGTGTGGGCCACCCACGCTGCGTGCCACCCGACCCCAACCCCCAGGAGGTGCTGGAGAACGTGACCGAGAACTTC  
 AACATGTGGAAGAACAACTGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACAGTCCCTGAAGCCCTGCGTGAAGCTGACCC  
 CCTGTGCGTGACCTGAACCTGACCGACCTGTGAACGCCACCAACACCACTCCACCAIATACCGTGGCGCGCGAGATCAAGAAT  
 GCTCCTTCAACATCACCACTCCATCCGCGACAAGATGCAGAGGATACGCCCTGTTCTACAAGCTGGACGTGGTGGCCATCGACAACAAC  
 ACCTCTACCGCTGATCAACTGCAACACCTCCGTGATCAACCGGCTGCCCCAAGGTGCTCTCGAGCCCATCCCCATCCACTACTGCAC  
 CCGCGCGGCTTCGCCATCTGAAGTGCAACGACAAGAAGTTCAACGGACCCGCCCCCTGCAAGAACGTGTCCACCGTSCAGTGCACCCACG  
 GCATCCGCCCCGTGGTGCCACCCAGCTGCTGTGAACGGCTCCCTGGCGGAGAGGAGGTGATCCGCTCCGAGAACTTCAACGACAAC  
 GCCAAGACCATCATCTGTGAGTGAACGAGTCCGTGGAGATCAACTGCACCCCGCCCCAACAAACACCCGAACTCCATCCACATCGGCC  
 CGGCCGCGCTTCTACGCCACCGCGGAGATCATCGCGGACATCCGCCAGGCCCATGCAACCTGTCCGCGCGCAAGTGAACAACACCCCTGA  
 AGCAGGTGGTGACCAAGCTGCGCGAGCAGTTCGACAACAAGACCATCTGTGTTCAACCCCTCTCCGCGCGGACCCCGAGATCGTGATGCAC  
 TCCTTCAACTGCGGCGGCGAGTCTTACTGCAACACCCACCGAGTGTTCAACTCCACCTGGACCGCACCTGGAACGACCCGAGAGGCAA  
 CATACCCCTGCCCTGCCGATCAAGCAGATCATCAACATGTGCGAGAGGTGGCAAGGCCATGTACGCCCCCCCCCATCCGCGGCCAGATCC  
 GCTGCTCCTCCAACATCACCGGCTGCTGTGACCCCGGACGCGGCAACAACGAGACCGAGATCTTCCGCCCCCGGCGCGCGACATGCGC  
 GACAACTGGCGCTCCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCTGGCGTGGCCCCCAACCAAGGCCAAGCCCGCGTGGTGCA  
 GCGGAGAAGCGCGCGTGGCATCGCGCCATGTTCTGGGCTTCTGGCGCGCGCGCTCCACCATGGGCGCGCCCTCCATGACCCCTGA  
 CCGTGCAGGCGCGCGAGCTGTGTCCGCGATCGTGCAGCAGCAGAACAACTGTGCGCGCCATCGAGGCCACGACGACCTGTGTCAGCTG  
 ACCGTGTGGGCGATCAAGCAGCTGCAGGCGCGCGTGTGCGCGTGGAGCGTACCTGCGCGACCGAGAGTCTGGGCACTGTGGGCTGCTC  
 CGGCAAGCTGATCTGCACCAACCGTGGTGGACGCTCTGCTGTTCCAACTGCTGGAGAGTCCCTGGACGAGATCTGGAACAACATGACCTGGATGG  
 AGTGGAGCGCGAGATCGACAACCTACACCGGCTGATCTAACCTGATCGAGGAGTCCAGAACCCAGCAGGAGAGAACGAGCAGGAGCTG  
 CTGGAGCTGGACAAGTGGGCTCCCTGTGGAACTGGTTCGACATCACCAACTGGCTGTGTATCATCAAGATCTTATCATGATGATGTTGGCGG  
 CCTGGTGGGCTTCGCGATCGTGTTCGCGCTGCTGTCCATCGTGAACCGGTGCGCGAGGCTACTCCCCCTGTCTTCCAGACCCGCTGC  
 CCGCCCCCGCGGCCCCGACCGGCGGCGGAGGAGGCGGCGGAGCGGACCGGACCGCTCCGCGCGCTGGTGAACGGCTTC  
 CTGGCCCTGATCTGGGACGACCTGGGCTCCCTGTGCTTCTCTACCAACCGCTGCGCGACCTGTGCTGATCGTGGCGCGCATCGTGGA  
 GCTGCTGGGCGCGCGGCTGGAGGCGCTGAAGTACTGTGGAACCTGCTGCAGTACTGGTCCCGAGGAGTGAAGAACTCCGCGGTGTC  
 TGCTGAACGCCACCGCCATCGCCGTGGCGGAGGCGACCGCGCTGATCGAGGTGGTGAAGCGCGCTTGCCTGCGGCGCATCTCTGCACATCCCC  
 CGCGCGATCCGCGAGGCGCTGGAGCGCGCGCTGCTGTA





79/178

Fig. 47A

## 2003 C.anc Env

MRVMGILRNCQQWMIWILGFWMIMCNVNVNLWVTYYGVVPVWKEAKTTLFCASDAKAYEREVHNWATHACVPTDPNPQEMVLENTENF  
 NMWKNMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATMGEKNCSEFNITTELDRDKKQVYALFYRLDIVPLNDNNSYRLINC  
 NTSAITQACPKVSFDPIPIHYCAPAGYAILKCNKTFNGTGPCNNVSTVQCTHGKIPVSTQLLNGSLAEIEIIIRSENLTDNAKTIIVHL  
 NESVEIVCTRPNNTTRKSIIRIGPGQTFYATGDIIGDIRQAHNCISEEKWKTQVRGEKLEHFPNKTIKFAPSSGGDLEITTHSFNCRGEF  
 FYCNTSRLFNSTYNSKNSTITLPCRICKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLTRDGGKNNETFRPGGDMRDNRSELYKYKV  
 VEIKPLGIAPTEAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLVQARQLLSGIVQQQSNLLRAIEAQHMLQLTVWGIKQLQTRV  
 LAIERYLKDQQLLGIWCSGKLICTTAVPWNSSWSNKSQEEIWDNMTWQWDREISNYTDTIYRLLEDSQOQEKNEQDALLADSWENLWNW  
 FDI TNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPNPRGPDRLGRIEEEGEQDRDRSIRLVSGFLALAWDDLRSLC  
 LFSYHRLRDLILIAARAVELLGRSSLRGLQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQICRAIRNIPRRIRQ  
 GFEAALL\$

Fig. 48A

## 2003 CON D Env

MRVRGIQRNYQHLLWRWGIMLLGMLMICSVAENLWVTYYGVVPVWKEATTLFCASDAKSYKTEAHNIWATHACVPTDPNPQEI ELENVTENF  
 NMWKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVKRNNTSNDTNEGEMKNCSEFNITTEIRDKKKQVHALFYKLDVVPIDDDNNSNT  
 SYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCKDKKFNGTGPKNVSTVQCTHGIRPVVSTQLLNGSLAEIEIIIRSENLTNNA  
 KIIIVQLNESVTINCTRPYNNTRQRTPIGPGQALYTRIKGDIRQAHNCISRAEWNKTLQOVAKKLGDLNKTIIIFKPSSGGDPEITTHSF  
 NCGGEFFYCNTSRLFNSTWNTKWNSTGKITLPCRICKQIINMWQGVKAMYAPPIEGLIKCSSNITGLLTRDGGANNSSHNETFRPGGDMR  
 DNWRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAIGLGAAGSTMGAASMTLVQARQLLSGIVQQQSNLLRAIEAQHLLQL  
 TVWGIKQLQARILAVERYLKDQQLGIWCSGKHICTTVPWNSSWSNKSLEIWNMTWMEWEREIDNYTGLIYSLIEESQOQEKNEQEL  
 LELEDKWLWNNWFSITQWLWYIKIFIMIVGGLIGLRIVFAVLSLVNRVRQGYSPLSFQTLTPAPRGPDPRPEGIEEGEQGRGRSIRLVNGF  
 SALIWDDLRNLCFLSYHRLRDLILIAARIVELLGRGWEALKYLNLLQYWIQELKNSAISLFDTTAIAVAEGTDRVIEIVQACRAILNIP  
 TRIRQGLERALL\$



80/178

Fig. 47B

## 2003 C. anc Env. seq. opt

ATCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGATCTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTTGGTGGG  
CAACCTGTGGGTGACCGTGCTACTACGGCGTGGCGTGGAGAGGCCAAGACCACCCCTGTCTTCTGCGCTCCGACGCCAAGGCCCTACGAGC  
GCGAGGTGCACAACGTTGGGCCACCCACGCTGCGTGGCCACCCAGACCCCAAGAGATGGTGTGGAGAACGTCGACCCGAGAACTTC  
AACAATGTGAAGAAGACATGTTGGACAGATGACAGGACATCATCTCCCTGTGGGACAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
CCTGTGCGTGACCCCTGAACGACCAAGCCCAACAGCCCAACCATGGGCGAGATGAAGAACTGCTCTTCAACATCACCCACCGAGC  
TGCGGCACAAGAAGCAGAGGTGTACGCCCTGTTCTACCGCTGGACATCTGTGCCCCCTGAACGACACAACACTCTTACCGCTGATCAACTGC  
AACACCTCCGCCATCACCCAGGCTGCCCAAGGTGCTTTCGACCCCATCCCATCTACTGCGCCCCCGCGGCTAGGCCATCCCTGAA  
GTGCAACAACAAGACCTTCAACGGCAACCGGCCCTGCAACAACGTGTCCACCGTGCACTGCAACCAAGCCATCAAGCCCGTGGTGTCCACCC  
AGCTGCTGTGAACGGCTCCCTGGCGGAGGAGGATCATCCGCTCCGAGAACCTGACCGACAAACGCAAGACCATCATCGTGCACCTG  
AACGAGTCCGTGGAGATCGTGTGACCCCGCCCAACAACAACACCCGCAAGTCCATCCGATCGGCCCGCGGACAGACCTTCTACGCCACCGG  
CGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGAGGAGAGTGAACAAGACCCCTGCAGCGCGTGGCGGAGAAAGCTGAAGG  
AGCACTTCCCCAACAAGACCATCAAGTTCGCCCTCTCCGCGGCGACCTGGAGATCACCACTCCATCACCTGCCCCCTGCGCATCAAGCAGATCATCAA  
TTCTACTGCAACACCTCCCGCTGTCAACTCACTCACTCAACTCCCAAGAACTCCACCATCACCTGCAAGTCCCAACATCACCGGCTGCTGCTGACCC  
CATGTGGCAGGGCGTGGCGGCCATGTACGCCCCCCCATCGCCGCAACATCACCTGCAAGTCCCAACATCACCGGCTGCTGCTGACCC  
GCGACGGCGGCAAGAACAACACCGAGACCTTCGCCCGGCGGCGGACATGCGCGCAACTGGCGCTCCGAGCTGTAAGTACAAGGTG  
GTGGAGATCAAGCCCCCTGGGCATCGCCCCACCGAGGCCAAGCCCGCTGGTGGAGCGCGAGAACGCGCCGCTGGGCATCGGCGCCGTGT  
CCTGGGCTTCTTGGGCGCCCGCGCTCCACCATGGGCGCCCTCCATCACCTGACCGTGCAGGCCCGCCAGCTGTCTCCGGCATCGTGC  
AGCAGAGTCCAACTGCTGGCGGCCATCGAGGCCAGCAGCATGTGCACTGACCGTGTGGGSCATCAAGCAGCTGCAGACCCCGCTG  
CTGGCCATCGAGCGCTACCTGAAGGACAGCAGCTGTGGGCTGTGGGCTGTCCGCAAGCTGATCTGACCAACCGGCTGCCCCTGGAA  
CTCTCTCTGGTCCAACAAGTCCAGGAGGAGATCTGGGACAACATGACCTGGATGCAGTGGACCCGCGAGATCTCCAACATACACCGACACCA  
TCTACCGCTGTGGAGGACTCCAGAACAGCAGAGAGAGAACGAGCAGGACCTGTGGCCCTGGACTCTTGGGAGAACCTGTGGAACTGG  
TTCGACATCACCAACTGGTGTGATCAATCAAGATCTTCAATCATGATCGTGGGCGCCCTGATCGGCCCTGGCATCATCTTCGCCGTGCTGTC  
CATCGTGAACCGCGTGGCAGGGCTACTCCCCCTGTCTTCCAGACCCCTGACCCCAACCCCGGCGCCCGACCCCTGGGCGCGCATCG  
AGGAGGAGGGCGGAGCAGGACCGGCTCCATCCGCTGTGGTTCGGCTTCTGGCCCTGGCTGGGACGACCTGCGCTCCCTGTGC  
CTGTTCTCTACCAACCGCTGCGGACTTCATCTGATCGCCCGCGCGCTGGAGCTGTGGGCGCTCTCCCTGCGCGGCTGCAGCG  
CGGCTGGAGGCCCTGAAGTACCTGGGCTCCCTGGTGCAGTACTGGGCGCTGGAGCTGAAGAAGTCCGCCATCTCCCTGTGGACACCATCG  
CCATCGCGCTGGCGAGGCGACCGCATCATCGAGCTGATCCAGCGCATCTGCGCGGCCATTCGCAACATCCCCCGCGCATCCGCGCAG  
GGCTTCGAGGGCGCCCTGCTGTA

**Fig. 48B**

2003 CON\_D Env.seq.opt

ATCGCGGTGGCGGCATCCAGCAACTACCAGCACTGTGGCGTGGGCGATCATGTGCTGGGCATGCTGATGATCTGCTCCGTGGCCGA  
GAACTGTGGTGACCGTGTAACGGGTGCCCGTGTGAAGGAGGCCACCACCACCTGTGTCTGCGCTCCGACGCCAAGTCTTACAAGA  
CCGAGGCCACAACATCTGGGCCACCCACGCTGCGTGCCACCGACCCCAAGGAGATCGAGCTGGAGAACGTGACCGAGAACTTC  
AATATGTGAAGAAACAATGTTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGACCAGTCCCTGAAGCCCTGCGTGAAGTGAACCCC  
CCTGTGCGTGACCTGAACTGCACCGACGTGAAGCGCAACAACCTCCAACGACACCAACGAGGCGAGATGAAGAACTGCTCTTCAACA  
TCACCAACGAGATCCGCGACAAGAAGACAGGTGCACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACGACAACAATCCAACACC  
TCCTACCGCTTGATCAACTGCAACACCTCCGCCATCACCCAGGCCCTGCCCAAGTGACCTTCGAGCCCATCCCATCTACTACTGCGCCCC  
CGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGAAGTCAACGGCACCGGCCCTCAAGAACGTGTCCACCGTGCAAGTGACACCCACGGCA  
TCCGCCCGTGGTGTCCACCGAGCTGCTGCTGAACGGTCCCTGGCCGAGGAGGATCATCTCCGTCCGAGAACCTTGACCAACAACGCC  
AAGATCATCATCGTGACGTGAACGAGTCCGTGACCATCAACTGCACCCGCCCTACAACAACCCGCCAGCGCACCCCATCGGCCCGG  
CCAGGCCCTGTACCAACCCGCATCAAGGGCGACATCCGCCAGGCCACTGCAACATCTCCCGCGCGAGTGAACAAGACCTGCAGCAGG  
TGGCCAAGAAGCTGGCGACCTGCTGAACAAGACCAACCATCATCTTCAAGCCCTCTTCCGGCGGACCCCGAGATCAACCACTCCTTC  
AACTGCGCGCGGAGTTCTTCTACTGCAACACCTCCCGCTGTTCAACTCCACTGGAACAACCAAGTGGAATCCACCGCAAGATCAC  
CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGCGAGGGCGTGGCAAGGCCATGTACGCCCCCCCATCGAGGGCTTGATCAAGTGCT  
CCTCCAACATCACCGCCTGCTGTGACCCGACGCGCGGCCAACAACTCCCAACAGACATTCGCCCCCGCGCGCGCGACATGCGG  
GACAACTGGCGCTCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTGGCGCTGCCCCACCCGCGCAAGCGCCGCTGGTGGGA  
GCGGAGAAGCGGCCATCGGCTGGCGGCCATGTTCTGGGCTTCTGGGCGCCCGCGCTCCACCATGGGCGCGCTCCATGACCCCTGA  
CCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGACGACGAGAACAACCTGCTGCGCGCCATCGAGGCCACGACGACCTGCTGCACTG  
ACCGTGTGGGCGATCAAGCAGCTGCAGGCCCGCATCCTGGCGGTGGAGCGTACCTGAAGGACGACGAGCTGCTGGGCATCTGGGGCTGCTC  
CGCAAGCACATCTGCACCACACCGTGCCCTGGAATCTCTCTGTTCCAAAGTCCCTGGACGAGATCTGGAACAACATGACCTGGATGG  
AGTGGAGCGCGAGATCGACAATACACCGCCTGATCTACTCCCTGATCGAGGAGTCCAGAACCCAGGAGAGAAGAACGACGAGGAGCTG  
CTGGAGCTGGACAAGTGGGCTTCCCTGTGGAATGTTCTCCATCACCCAGTGGCTGPGTACATCAAGATCTTCAATCATGATCGTGGCGG  
CCTGATCGGCCCTGCGCATCGTGTTCGCCGTGCTTCCCTGGTGAACCGGTGCGCCAGGGCTACTCCCGCTGTCTTCCAGACCCCTGCTGC  
CCGCCCCCGGCCCGACCCCGAGGGCATCGAGGAGGCGGCGAGAGGGCCGCGCGCTCCATCCGCCCTGGTGAACGCGCTTC  
TCCGCCCTGATCTGGGACGACCTGCGCAACCTGTGCCCTGTTCTCTACCAACCGCTGCGCGACCTGATCCTGATCGCCCGCGCATCGTGGA  
GCTGCTGGGCGCGCGGCTGGAGGCCCTGAAGTACCTGTGGAACCTGTGCACTGATCCAGGAGCTGAAGAACTCCGCCATCTCCC  
TGTTGACACCAACCGCCATCGCCGTGGCGGAGGACCGGACCGGTGATCGAGATCGTGACGCGCGCTGCCGCGCCATCCTGAACATCCCC  
ACCGCATCCGCCAGGGCCTGGAGCGCGCCCTGCTGTAA

82/178

Fig. 49A

2003 CON F1 Env

MRVRGMQRN̄WQHGLGKWGLLEFLGILIIICNAADNLWVTYYGVVPVWKEATTLFCASDAKSYEKEVHNWVWATHACVPTDPNPQEVVLENVTEF  
 DMWKNM̄VQMHEDIIISLWDQSLKPCVKLTPLCVTLNCTDVNATNDTNDNKTGAIQNCSEFNMTTEVRDKKLVHALFYKLDIVPISNNNSK  
 YRLINCNTSTITQACPKVSDPIPIHYCAPAGYAILKCNDKRFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEDIIIRSONISDNAK  
 TIIVHLNESVQINCTRPNNNTRKSIHLGPGQAFYATGEIIGDIRKAHCNISGTQWNKTLEQVAKLKSHPNKTIKFNSSSGGDLEITMHSF  
 NCRGEFFYCNTSGLENDTGSNGTITLPCRIKQIVNMWQEVGRAMYAAPIAGNITCNSNITGLLLTRDGGQNNTETFRPGGGMKDNWRSELY  
 KYKVVEIEPLGVAPTKAKRQVVKRERRAVGIGAVFLGELGAGSTMGAASITLTVOARQLLSGIVQQQNLLRAIEAQHLLQLTVWGKQL  
 QARVLAVERYLKDQQLLGLWGCSGKLICTTNVPWNSSWSNKSQDEIWNMTWMEWEKEISNYSNIIYRLIEESQNOQKEQELLALDKWAS  
 LWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNVRKGYSPLSQTLIPSPREPDRPEGIEEGGEGQKDRSVRLVNGFLALVWDDL  
 RNLCLFSYRHLRDFILIAARIVDRGLRGRGWEALKYLGNTLTQYWSQELKNSAISLLNTTAIVVAEGTDRVIEALQAGRAVLNIPRRIRQGLE  
 RALL\$

Fig. 50A

2003 CON F2 Env

MRVREMQRN̄WQHGLGKWGLLEFLGILIIICNAADNLWVTYYGVVPVWKEATTLFCASDAKAYEREVHNWVWATYACVPTDPSQELVLGNVTENE  
 NMWKNM̄VDQMHEDIIISLWDQSLKPCVKLTPLCVTLNCTDVNVTINTNVTLGEIKNCSEFNITTEIKDKKKKEYALFYRLDVVPINN̄SIVYR  
 LISCNTSTVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGLCRNVSTVQCTHGIRPVSTQLLNGSLAEEDIIIRSENISDNTKTI  
 IVQFNRSVEINCTRPNNNTRKSIRIGPGRAFYATGDIIGDIRKAYCNINRTLWNETLKKVAEEFKNHFNITVTFNPSSGGDLEITTHSFNCR  
 GEFFYCNTSDLEFNTEVNNTKITITLPCRIRQFVNMMQVRGRAMYAPPIAGQIQCNISNITGLLLTRDGGKNGSETLRPGGDMRDNRSELYK  
 YKVVKIEPLGVAPTKAKRQVQREKRAVGIGAVLGLGAGSTMGAASITLTVOARQLLSGIVQQQNLLKAIQAQHLLQLTVWGKQLQ  
 ARILAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMQWEKEISNYTDTIYRLIEDAQNOQKEQDLLALDKWDNL  
 WSWFTITNWLWYIKIFIMIVGGLIGLRIVFAVLSVNVNRQGYSPLSQTLIPNPRGPERPGGIEEGGEGQDRDRSIRLVSGFLALAWDDL  
 SLCLFSYRHLRDFILIAARTVDMGLKRGWEALKYLWNLPOYWGQELKNSAISLLDTTAIAVAEGTDRIIEVLQAGRAVLNIPRRIRQGFER  
 ALL\$

83/178

Fig. 49B

## 2003 con F1 Env. seq. opt

ATGGCGGTGCGGGCATGCAGCGCAACTGGCAGCACCTGGGCAAGTGGGGCTGTGTCTCTGGGCATCCTGATCATCTGCAACGCCGCCGA  
 GAACCTGTGGGTGACCGTGTACTACGGCGTGCCTGTGGAGGAGGCCACCAACCACTGTCTGGCCCTCCGACGCCAAGTCTCTACGAGA  
 AGGAGGTGCACAACAGTGTGGGCCACCCACGCTGCGTGCACACCCCAACCCAGGAGGTGGTGGAGAACGTGACCGAGAACTTC  
 GACATGTGGAAGAACAACATGTTGGAGCAGATGCACACCGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCCCTGAACCTGCACCGACGTGAACGCCCAACCAACGACACCAACGACCGGCCCATCCAGAACTGCTCCTTCA  
 ACATGACCACCGAGGTGCGGACAAAGAGTGAAGGTGACGCCCTGTTTACAAGCTGGACATCGTGGCCATCTCCAACAACAACCTCAAG  
 TACCGCTGATCAACTGCAACACCTCCACCATCAACCCAGGCTGCCCAAGGTGTCTGGACCCCATCCCATCTCACTACTGGCCCCCGC  
 CGGTACGCCATCCTGAAGTGCAACGACAGCGCTTCAACGGCACCCGCCCTTGAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCA  
 AGCCCGTGGTGTCCACCCAGCTGTCTGAACGGCTCCCTGGCCGAGGAGACATCATCCGCTCCCAAGAACATCTCCGACAAACGCCAAG  
 ACCATCATCGTGCACTGAACGAGTCCGTGCAGATCAACTGCACCCGCCCAACAACAACCCGAAGTCCATCCACCTGGGCCCGGCCA  
 GGCCTTACGCCACCCGGCGAGATCATCGCGGACATCCGCAAGGCCCACTGCAACATCTCCGGCACCCAGTGGAAACAAGACCTTGGAGCAGG  
 TGAAGGCCAAGTGAAGTCCCACTTCCCAACAAGACCATCAAGTTCAACTCTCTCCGGGGACCTGGAGATCACCATGCACTCCTTC  
 AACTGCCGCGCGAGTTCTTCTACTGCAACACTCCGGGCTGTTCAACGACACCGGCTCCAACGGCACCATCACCTGCCCTGCCGTCATCAA  
 GCAGATCGTGAACATGTGGCAGGAGTGGGCCGCCCATGTACGCCGCCCATCGCCGCAACATCACCTGCAACTCCAACATCACCGGCC  
 TGCTGTGACCCGGCAGCGGCCAGAACAAACCGAGACCTTCCGCCCGCGCGCAACATGAAGGACAACTGGCGCTCCGAGCTGTAC  
 AAGTACAAGGTGGTGGAGATCGAGCCCTGGCGTGGCCCCACCAAGGCCAAGCGCCAGGTGGTGAAGCCGAGCGCCCGCGCTGGGCAT  
 CGCGCGCTGTCTGGGCTTCTGGGCTCCACCATGGGCGCGCTCCACCATGGGCGCGCTCCATCACCTGACCGTGCAGGCCCGCGCTGTGT  
 CCGGCATCGTGACAGCAGAACAACTGTGTGGCGCATCGAGGCCAGCAGACCTGTGCAAGTGAAGTGGGGCTGCTCCGGCAAGCTGATGCAACCAA  
 CAGGCCCGCTGTGGCGCTGAGCGCTACCTGAAGGACCAAGTCCAGGACGAGATCTGGAACAACATGACCTGGATGGAGTGGGAGAGAGATCTCCA  
 ACTCCAACATCATCTACCGCTGATCGAGGAGTCCAGAACCCAGCAGGAGAGAACAGCAGGAGCTGTGGCCCTGGACAAAGTGGGCTCC  
 CTGTGGAACCTGGTTCGACATCTCCAACCTGGCTGTGGTACATCAAGATCTTCAATGATCGTGGCGGCTGATCGGCTTGGCATCGTGT  
 CGCGTGTGTCCATCGTGAACCGCGTGCAGAGGCTACTCCCCCTGTCCCTGCAGACCTGATCCCCCTCCCGCGGAGCCGACCGCC  
 CCGAGGGCATCGAGAGGCGCGCGGAGCAGGCAAGACCGCTCCGTGGCTGGTGAACGGCTTCTGGCCCTGGTGTGGACGACCTG  
 CGCAACCTGTGCTTCTCTACCGCCACCTGCGCGACTTCACTCTGATCGCGCGCGCATCGTGGACCGCGGCTGCGCGCGGCTGGGA  
 GGCCCTGAAGTACCTGGGCAACCTGACCCAGTACTGGTCCAGGAGTGAAGAACTCCGCCATCTCCTGTGAACACCAACCGCCATCGTGG  
 TGGCCGAGGACCGACCGCGTGTGAGAGGCCCTGCAGCGCGCGCGCGCTGTGTAACATCCCCCGCGCATCCCGCAGGGCCTGGAG  
 CGCGCCCTGCTGTAA

84/178

Fig. 50B

2003 CON F2 Env. seq. opt

ATGGCGGTGGCGAGATGCAGGCAACTGGCAGCACCTGGGCAAGTGGGCGCTGCTGTCTCTGGGCATCCTGATCATCTGCAACGCCGCCGGA  
CAACCTGTGGGTGACCGGTACTACGGCGTGGGAGGAGGCCACCAACCTGTCTGGCGCTCCGACGCCAAGGCTACGAGC  
GCGAGGTGCACAACGTTGGGCCACCTACGCTGCGTGCCACCGACCCCTCCCGCAGGAGCTGGTGGGCAACCTGACCGGAACTTC  
AACATGTGAAGAACAACATGTTGGACCATGTCAGGAGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
CCTGTGCGTGACCTGAACCTGCACCGACGTGAACGTGACCATCAACACCAACCAACCTGACCTGGGCGAGATCAAGAACTGCTCCTTCAACA  
TCACCAACCGAGATCAAGGACAAGAAGAGTAGCCCTGTTCTACCGCTGGACGTGGTGCCCATCAACAACTCCATCGTGTACCGC  
CTGATCTCCTGCAACACCTCCACCGTGACCCAGGCTGCCCAAGGTGCTCTCGAGCCCATCCCATCTCACTACTGCGCCCCCGCGGCTT  
CGCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCTGTGCCGCAACGTGTCCACCGTGCACTGACCCACGGCATCCGCCCG  
TGGTGCCACCCAGCTGCTGTAACGGCTCCCTGGCCGAGGAGACATCATCCGCTCCGAGAACATCTCCGACAAACACCAAGACCATC  
ATCGTGCAGTTCAACCGCTCCGTGGAGATCAACTGCACCCCGCCCAACAACAACCCGCAAGTCCATCCGCATCGGCCCGCGCGCTT  
CTAGCCCAACCGCGACATCATCGCGACATCCGCAAGGCTACTGCAACATCAACCGCACCTGTGGAAAGAGACCTTGAAGAGGTGGCG  
AGGAGTTCAAGAACCATTCAACATCACCGTGACCTTCAACCCCTCCTCCGGCGGCGACCTGGAGATCAACACCACTCCTTCAACTGCCGC  
GGCGAGTTCTTCTACTGCAACACCTCCGACCTGTTCAACAACAACCGAGTGAACAACAACCAAGACCATCACCTGCCCTGCCGATCCGCCA  
GTTCTGTGAACATGTGGCAGCGGTGGCGCGGCGCATGTACGCCCGCCCATCGCGGCGAGATCCAGTGCAACTCCAACATCACCGGCTGC  
TGCTGACCCGCGACGGCAAGAACGGCTCCGAGACCTCGCGCGCGGCGACATGCGGACAACTGGCGCTCCGAGCTGTACAAG  
TACAAGGTGTTGAAGATCGAGCCCTTGGCGGTGGCGCGCGGCGCGAGGTGGTGCGAGCGGAGAGCGCGCTGGGATCGG  
CGCGTGCTGCTGGGCTTCTGGCGCGCGCGGCTCCACCATGGCGCGCGCTCCATCACCTGACCTGACCGTGCAGGCGCGCGCTGGGATCGG  
GCATCGTGACGACGATCCAACTGCTGAAGGCCATCGAGGCCAGCACCTGCTGCAGCTGACCGTGTGGGCAATCAAGCAGCTGCAG  
GCCCGCATCCTGGCGGTGGAGCGCTACCTGAAGGACCAGCAGCTGTGGGCTGTGGGCTGCTCCGGCAAGCTGATCGCACCAACCAAGT  
GCCCTGGAACCTCCTGTTCCAAAGTCCAGGACGAGATCTGGACAACATGACCTGGATGCGAGTGGGAGAGGAGATCTCCAACATACA  
CCGACACCATCTACCGCTGATCGAGGACGCCAGAACCCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGACAACCTG  
TGGTCTGTTTACCATCAACATGCTGTGTATCATCAAGATCTTCAATCATGATCGTGGGCGGCTGATCGGCCCTGGCATCGTGTTCGC  
CGTGTGTCGTTGTAACCGGTGCGCCAGGCTACTCCCGCTGTCCCTGCAGACCTGATCCCCAACCCCGCGCGCGCGGAGCGCGCG  
GCGGCATCGAGGAGGCGGCGAGCAGGACCGGACCGCTCCATCCGCTGGTGTCCGCTTCTCTGGCCCTGGGACGACCTGCGC  
TCCCTGTGCTGTTCTCTACCGCACCTGCGGACTTCACTCTGATCGCGCGCGGACCGCTGGACATGGCCCTGAAGCGCGGCTGGGAGGC  
CCTGAAGTACCTGTGGAACCTGCCCCAGTACTGGGCGCAGGAGTGAAGAACTCCGCGCATCTCCCTGTGTGACACCAACCGCATCGCGCTGG  
CCGAGGGCACCGACCGCATCATCGAGGTGCTGACGCGCGCGCGGCTGCTGCACATCCCCCGCGCGCATCCCGCCAGGGCTTCGAGCGC  
GCCCTGCTGTAA

85/178

Fig. 51A

## 2003 CON G Env

MRVKGIQRNWQHLLKWTLLILGLVICSASNLLWVTYYGVVWEDADTTLCASDAKAYSTERHNWVWATHACVPTDPNPQEITLENVTFENF  
 NMWKNMVEQMHEDEIISLWDESLKPCVKLTPLCVTLNCTDVNTNNTNNTKKEIKNCSEFNITTEIRDKKKKEYALFYRLDVVPINDNGNSS  
 IYRLINCNVSTIKQACPKVTFDPIPIHYCAPAGFAILKCRDKKENGTPCKNVSTVQCTHGKPKPVSTQLLINGSLAEEEEIIIRSENITDNT  
 KVIIIVQLNETIEINCTRPNNTRKSIIRIGPGQAFYATGDIIGDIRQAHCVSRKWNEMLQVKAQLKKIFNKISITFNSSSGGDLITTHSF  
 NCRGEFFCYNTSGLENNLSNSTNSTITLPCIKIQIVRMWQVQAMYPPIAGNITCRSNITGLLLTRDGGNNNTETFRPGGDMRDNRWS  
 ELYKIKIVKIKPLGVAPTRARRRVVEREKRAVGLGAVLLGFLGAAGSTMGAASITLTVOVRQLLSGIVQQSNLLRAIEAQHLLQLTVWGI  
 KQLOARVLAVERYLKDQQLLGWCSGKLICTTNVPWNTSWSNKSNEIWDNMTWIEWEREISNYTQQIYSLIEESQKQKNEQDLLALDK  
 WASLWNWFDTIKWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLTHHQREPDPERIEEGGEGQDKDRSIRLVSGFLALAW  
 DDLRSLCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGLYLWNLLLYWGQELKNSAINLLDTIAIAVANWTDRIEVAQACRAILN  
 IPRRIRQGLERALL\$

Fig. 52A

## 2003 CON H Env

TRVMEQQRNYPSSLWRWGTLILGMLLICSAGNLWVTYYGVVWKEAKTTLCASDAKAYETEKHNWVWATHACVPTDPNPQEMVLENVTENF  
 NMWENDMVEQMHTDIIISLWDSQSLKPCVKLTPLCVTLDCSNVNTTNAATNSRFNMQEELTNCSEFNVTTVIRDKQKQKHALFYRLDVVPIDDNNS  
 YQYRLINCNTSVITQACPKVSEFPIPIHYCAPAGFAILKCNKKTFTNGTGPCTNVSTVQCTHGIRPVWSTQLLINGSLAEEQVVIIRSKNISDN  
 TKNIIIVQLNKPVEITCTRPNNTRKSIHLGPGQAFYATGDIIGDIRQAHCVSRKWNEMLQVKAQLKKIFNKISITFNSSSGGDLITTHSF  
 SENCERGEFFCYNTSGLENNLSNSTNSTITLPCIKIQIVRMWQVQAMYPPIAGNITCRSNITGLLLTRDGGNNNTETFRPGGDMRDNRWS  
 NWRSELYKYKVVKIEPLGVAPTRARRRVVEREKRAVGMGAFFLGLGAAGSTMGAASITLTVOVRQLLSGIVQQSNLLRAIEAQHLLQLTVWGI  
 VWGIKQLOARVLAVERYLKDQQLLGWCSGKLICTTNVPWNTSWSNKSNEIWDNMTWIEWEREISNYTQQIYSLIEESQKQKNEQDLLALDK  
 ALDKWASLWNWFSTINWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDPRPEGIEEGGEGQDRDRSVRLVNGFL  
 PLVWDDLRSLCLFSYRLLRDLLIIVVRTVELLGRRRGREALKYLWNLLQYWGQELKNSAINLLNTTIAIAVAEGTDRIEIVQRAWRAILHIPR  
 RIRQGFERTLL\$

86/178

Fig. 51B

2003 CON G Env. seq. opt  
ATGCGGTGAAGGGCATCCAGCGCAACTGGCAGCACCTGTGGAAAGTGGGGCACCCCTGATCCTGGGCTGGTGATCATCTGCTCCGGCTCCAA  
CAACCTGTGGGTGACCGTGTACTACGGCGTGCCGTGTGGAGGACGCCGACACCACTCTGTTCTGGCTCCGACGCCAAGGCTACTCCA  
CCGAGCGCCACAACGTGTGGCCACCCACGCTGCGTGCCACCGACCCCAACCCACGAGATCACCTGGAGAAAGTGACCGGAACTTC  
AACATGTGAAGAACAACATGGTGGAGCAGATGACGAGGACATCATCTCCTGTGGACGAGTCCCTGAAGCCTGCGTGAAGCTGACCCC  
CCTGTGCGTGACCCCTGAACCTGCACCGACGTGAACGTGACCAACAACACCAACAAGAGAGATCAAGAACTGCTCTTCAACA  
TCACCAACCGAGATCCGCGACAAGAAGAGGAGTACGCCCTGTTCTACCGCTGGACGTGGTCCCATCAACGACAAACGGCAACTCCTCC  
ATCTACCGCTGATCAACTGCAACGTGTCCACCATCAAGCAGGCTGCCCAAGGTGACCTTCGACCCCATCCCATCTACTGCGCCCC  
CGCCGCTTCGCCATCCTGAAGTCCCGGACAAAGATTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAAGTGACCCACGGCA  
TCAAGCCCCGTGTCCACCCAGCTGTCTGAACGGCTCCCTGGCCGAGGAGGAGATCATATCCGCTCCGAGAAACATCACCGACAAACAC  
AAGTGATCATCTGCAGCTGAACGAGACCATCGAGATCAACTGCACCCCGCECAACAACAACACCCGCAAGTCCATCCGATCGGCCCGG  
CCAGGCTTCTACGCCACCGCGGACATCATCGCGACATCCGCCAGGCCACTGCAACGTGTCCCGCACCAAGTGGAAACGAGATGCTGCAGA  
AGTGAAAGGCCCCAGCTGAAGAAGATCTTCAACAAGTCCATCACTTCAACTCCTCCGCGGCGACCTGGAGATCACCACTCCCTTC  
AATGCCGCGGAGTCTTCTACTGCAACACCTCCGCTGTCAACAACCTCCCTGTGAATCCCAACTCCACCATCACCTGCCCCCTG  
CAAGATCAAGCAGATCGTGGCATGTGGCAGCGGTGGCCAGGCTGTACGCCGCCCATCGCCGGAACATCACCTGCCCTCCAAACA  
TCACCGCTGTGTACCCCGGACCGCGGCAACAACAACCGAGACCTTCGCCCGCGGCGGCGACATGCGGACAACTGGCGCTCC  
GAGCTACAAGTACAAGATCGTGAAGATCAAGCCCCTGGGCTGGCCCCACCCGCGCCGCGGTGGTGGAGCGCGAGAGCGGC  
CGTGGGCTGGGCGCGTGTGTGGCTTCTTGGGCGCGCGGCTCCACCATGGGCGCGCTCCATCACCTGACCGTGCAAGTGGGCATC  
AGCTGCTGTCCGGCATCGTGACAGCAGTCCAACCTGTGCGCGCATGAGGCCACAGCACCTGTGGCTGTGGGCTGTCCGGCAAGCTGATCTG  
AAGCAGCTGCAGGCCCGCTGTGGCGTGGAGCGTACCTGAAGGACCAAGCAGTGTGGGCTGTGGGCTGTCCGGCAAGCTGATCTG  
CACCACCAACGTGCCCTGGAACACCTCCTGTGTCACAACAGTCTTACAACGAGATCTGGGACAAACATGACCTGGATCGAGTGGAGCGCGAGA  
TCTCCAACCTACACCCAGCAGATCTACTCCTGATCGAGGAGTCCCAAGAACCAAGCAGGAGAGAACGAGCAGGACCTGTGGCCCCCTGGACAAG  
TGGGCTCCCTGTGGAACCTGATCGACATCAACCAAGTGGTGTGGTACATCAAGATCTTCAATGATCGTGGGCGGCTGATCGGCTGCG  
CATCGTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGGCCAGGCTACTCCCCCTGTCTTCCAGACCTTGAACCAACAGCGCGAGC  
CCGACCGCCCCGAGCGCATCGAGAGGCGCGGCGAGCAGGACCGCTCCATCCGCTGGTGTCCGCTTCTTGGCTTCTTGGCTGGCTGG  
GACGACCTGGCTCCCTGTGCTTCTTCTTACCACCGCTGGCGACTTCATCCTGATCGCGCGCGCACCGTGGAGTGTGGGCTGGCTC  
CTCCCTGAAGGCTTGGCTGGGCTGAGTACCTGTGAACCTGTGTGAGGCTGTGAGGCTGAAGAACTCCGCCA  
TCAACCTGCTGGACACCATCGCCGTGGCCAACTGGACCGGCTGATCGAGGTGGCCAGCGCGCTGCCGCGCATCTCTGAAC  
ATCCCCCGCGCATCCGCCAGGGCTTGAGCGGCCCTGTCTGTA



87/178

Fig. 52B

2003 CON H Env.seq.opt

ACCCGGTGATGGAGACCCAGCGCAACTACCCCTCCCTGTGGCGTGGGACCCCTGATCCTGGGCATGCTGCTGATCTGCTCCGCCGCCG  
 CAACCTGTGGGTGACCGTGTACTACGGCTGCCCGTGGAAAGGAGGCCAAGACCAACCTGTTCTGCGCTCCGACGCCAAGGCTACGAGA  
 CCGAGAAGCACAACTGTGGGCCACCCACGCTGCCGTGCCACCCCAACCCACAGGAGTGGTGGAGAACTGACCGAGAATCTC  
 AACATGTGGAGAACGACATGTTGGAGCAGATGCACACCGACATCATCTCCCTGTGGACCAAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGGTGACCCCTGGACTGCTCCAACTGAACACCAACCGCCCACTCCCGCTTCAACATGCAGGAGGAGCTGACCAACTGCTCCT  
 TCAACGTGACCACCGTGTCCCGGACAAAGCAGAGAGTGCACGCCCTGTTCTACCGCTGGACGTGGTCCCATCGACGACAACTCC  
 TACCAGTACCGCTGATCAACTGCAACACCTCCGTGATCACCCAGGCTGCCCAAGGTGCTCTCGAGCCCATCCCATCTACTGCGC  
 CCGCGCGCTTCGCCATCCTGAAGTGCAACAACAGACCTTCAACGGCACCGGCCCTGCACCAACGTGTCCACCGTGCAGTGCACCCACG  
 GCATCCGCCCGTGTCCACCCAGCTGCTGAACGGCTCCCTGGCCGAGGAGCAGGTGATCATCCGCTCCAAGAACATCTCCGACAA  
 ACCAAGAACATCATCGTGACGTGAACAAGCCGTGGAGATCACTGACCCGCCCAACAACAACACCCGCAAGTCCATCCACCTGGGCC  
 CGCCAGGCTTCTACGCCACCGCGGACATCATCGCGGACATCCGCCAGGCCACTGCAACATCTCCGGCAAGAAGTGAACAAGACCTGC  
 ACCAGGTGGTGAACCCAGCTGGGCAAGTACTTCGACAAACCGACCATCATCTTCAAGCCCACTCCGGCGGCGACATGGAGTGACCCAC  
 TCCTTCAACTGCCCGCGGAGTCTTCTACTGCAACCTCCGGCTGTCAACTCCTCTGGACCACTCCACCAACGACCAAGAACAT  
 CATACCTGCCCTGCCATCAAGCAGATCGTGAACATGTGGCAGCGGTGGCCAGGCCATGTACGCCCTCCCTCCATCAAGGCAACATCA  
 CTTGCGTGTCCAACATCACCGGCTGATCCTGACCTTCGACGAGGCAACAACACCGTGACCTTCCGCCCGCGCGGCGGACATGCCGAC  
 AACTGGCGTCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTGGCGGTGGCCCAACCGAGGCCCGCGCTGCTGCTGAGCG  
 CGAGAAGCGCGCTGGGCAATGGCGCTTCTTCTGGCTTCTGGCGCGCGCGCTCCACCATGGCGCGCGCTCCATCACTGACCG  
 TGCAGGCGCGCGAGCTGTCCGGCATCGTGAGCAGCAGTCCAACTGTGCGCGCCATCCAGGCCAGCAGCATGTGCAGCTGACC  
 GTGTGGGCGCATCAAGCAGCTGCAGGCCCGGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAAGCTGCTGGGCTGTGGGCTGCTCCGG  
 CAAGCTGATCTGCACCAACAGTGCCCTGGAACCTCCTCTGGTCCAACAGTCCCTGGACGAGATCTGGGACAACTGACCTGGATGGAGT  
 GGGACAAGCAGATCAACAATACACGAGGAGATCTACCGCTGTGGAGGTGTCCAGACCCAGCAGGAGAAACGAGCAGGACCTGTG  
 GCCCTGGACAAGTGGGCTCCCTGTGGAACTGGTCTCCATCAACCACTGGTGTGATCAAGATCTTATCATGATCGTGGGCGGCT  
 GATCGGCTGCCATCATCTTCGCCGTGTGTCCATCGTGAACCGCTGCGCCAGGCTACTCCCTGTCTTCCAGACCTGATCCCCA  
 ACCCCGCGGCCCGACCGCCCGAGGCGATCGAGGAGGAGGCGGAGCAGGACCGGACCGCTCCGTGCCCTGGTGAACGGCTTCCCTG  
 CCGCTGGTGTGGACGACCTGCGCTCCCTGTGCCCTGTCTCCCTACCGCTGTGCGGACCTGCTGTGATCGTGGTGGCACCGTGGAGCT  
 GCTGGGCGCGCGGCGGAGGCGCTGAAGTACCTGTGAACTGTGAGTACTGGGCGCAGGAGTGAAGAACTCCGCCATCAACCTGC  
 TGAACACACCGCCATCGCCGTGGCGGAGGCAACCGACCGCATCATCGAGATCGTGACGCGCGCTGGCGCGCCATCCTGCACATCCCCCGC  
 CGCATCCCGCAGGGCTTCGAGCGCACCCCTGCTGTAA



88/178

Fig. 53A

2003 CON 01 AE Env

MRVKETQMNPNLWKWGTLILGLVICSASDNLWVTVYGVVWRDADTTLCASDAKAHETEVEHNVWATHACVPTDPNPQEIHLNVTEF  
 NMWKNMVEQMDEIISLWDQSLKPCVKLTPLCVTLNCTNANLTNVNNITVSNIIIGNITNEVRNCSFNMTELDRKKQKVHALFYKLDIVQ  
 IEDNNSYRLINCNTSVIKQACPISFDPIPIHYCTPAGYAILKCNDRKFNCTGPKCNVSSVQCTHGKIPVSTQLLLNGSLAEEIIIRSEN  
 LTNNAKTIIVHLNKSVEINCTRPSNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEVLKQVTEKLKEHFNNKTIIFQPPSGGDL  
 ITMHENCERGEFFCYNTTKLFNNTCIGNETMEGCNGTIIILPKIKQIINMWQAGQAMYPPIISGRINCVSNITGILLTRDGGANNTNETFR  
 PGGNIKDNWRSELYKYKVQIEPLGIAPTRAKRRVVEREKRAVGIGAMIFGLGAGSTMGAASITLTVOARQLLSGIVQQQSNLLRAIEA  
 QOHLQLTVWGIKQLOARVLAVERYLKDKQFLGLWCGSGKIICTTAVPNSTWSNRSEFEEIWNMTWIEWEREISNYTNQIYEILTESQNQQ  
 DRNEKDLLELDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTPTHHQREPDPRPERIEEGGGEQGRDRS  
 VRLVSGFLALAWDDLRSCLFSYHRLRDFILIAARTVELLGHSSSLKGLRRGWEGLKYLGNLLLYWGQELKISAILLDDATAIAVAGWTDREVI  
 EVAQGAWRAILHIPRRIRQGLERALL\$

Fig. 54A

2003 CON 02 AG Env

MRVMGIQKNYPILWRGMIIFWIMIICNAENLWVTVYGVVWRDAETTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIHLNVTEFEN  
 MWKNMVEQMDEIISLWDQSLKPCVKLTPLCVTLNCTNANLTNVNNITNSNTNNAEIKNCSFNMTELDRKKQKVYALFYRLDVVQINKNNSQYR  
 LINCNTSAITQACPVSFEPIPIHYCAPAGFAILKCNDRKFNCTGPKCNVSTVQCTHGKIPVSTQLLLNGSLAEEIIIRSENITNNAKTI  
 IVQLVKPVKINCTRPNNTKRSVRIGPGQTFYATGDIIGDIRQAHCVSRTKWNNTLQQVATQLRKYFNKTIIFANPSGGDLITTHSFNCG  
 GEEFYCNTSELFNSTWNSTWNTTEKICITLQCRIKQIIVNMWQVQAMYPPIQGVIRCESNITGILLTRDGGNNSTNETFRPGGDMRDNW  
 RSELYKYKVQIEPLGVAPTRAKRRVVEREKRAVGIGLGAAGSTMGAASITLTVOARQLLSGIVQQQSNLLRAIEAQHLLKLTIVW  
 GIKQLOARVLAERYLKDQQLGIWCGSGKLICTTVPWNSSWNSKTYNDIWNMTWLQWDKEISNYTDIIYNLIEESQONQEKNEQDILLAL  
 DKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVEAVLTIINVRQGYSPLSFQTLTHHQREPDPRPERIEEGGGEQDRDRSVRLVSGFLAL  
 AWDDLRSCLFSYHRLRDFVLIARTVELLGHSSSLKGLRLGWEALKYLGNLLSYWGQELKNSAINLLDTIAIAVANWTDREVIIEIGQAGRAI  
 LNIIPRRIRQGLERALL\$

89/178

Fig. 53B

2003 CON 01 AE Env.seq.opt

ATGCGCGTGAAGGAGACCCAGATGAACCTGGCCCAACCTGTGGAAGTGGGCAACCTGATCCTGGGCCCTGGTGATCATCTGCTCCGCCCTCCGA  
CAACCTGTGGGTGACCGTGTACGGCGTGTGGCGGACGCCGACACCAACCTGTCTGGCCCTCCGACGCCAAGGCCCAACGAGA  
CCGAGGTGCACAACCGTGTGGGCCACCCACGCTGCGTGGCCACCCGACCCCAACCCAGAGATCCACCTGGAGAACGTGACCGAGAACTTC  
AACATGTGAAGAACAACATGTTGGAGCAGATGCAGGAGGACGTGATCTCCCTGTGGACCACTCCCTGAAGCCCTGGCGTGAAGCTGACCCC  
CCTGTGCGTGAACCTGAACCTGCAACACCTGCAACACCTCGGTGATCAAGCAGGCCCTGCCCAAGATCTCCTTCGACCCCATCCCCAT  
AGGTGCGCAACTGCTCTCAACATGACCAACCGGCTGCGGACCAAGAGAGGTGCACGCCCTGTTCTACAGCTGGACATCGTGCAAG  
ATCGAGGACAACAACCTCTACCGCTGATCAACCTCGGTGATCAAGCAGGCCCTGCCCAAGATCTCCTTCGACCCCATCCCCAT  
CCACTACTGCACCCCGCGCTACGCCATCCTGAAGTGCAACGACAAGAACTTCAACGGCACCGGCCCTGCAAGAACGTGTCTCCGTGC  
AGTGACCCACCGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGAGATCATCATCCGCTCCGAGAAC  
CTGACCAACAACGCCAAGACCATCATCGTGCACTTGAACAAGTCCGTGGAGATCAACTGCACCCGCCCTCCAAACAACACCCGACCTCCAT  
CACCATCGGCCCGCCAGGTGTTCTACCGCACCGCGACATCATCGCGACATCCGCAAGGCCCTACTGCGAGATCAACGGCACCAAGTGGA  
ACGAGGTGCTGAAGCAGGTGACCGAGAGCTGAAGGAGCACTTCAACAACAAGACCATCATCTCCAGCCCCCTCCGGCGGACCTGGAG  
ATCACCATGCACCACTTCAACTGCCCGCGGAGTCTTCTACTGCAACACCAACCAAGCTGTTCACAACAACACCTGCATCGGCAACGAGACCAT  
GGAGGGTGCACCGCACCATCATCTCGCTGCTCAAGATCAAGCAGATCATCAACATGTGGCAGGGCGCGCCAGGCCATGTACGCCCCCC  
CCATCTCCGGCCGATCAACTGCGTGTCCAAACATCAACGGCATCCTGTGACCCGACGGCGGCCCAACAACAACGAGACCTTCCGC  
CCCGCGCGGCAACATCAAGACAACCTGGCGCTCCGAGCTGTACAAGTACAAGGTGTGACAGATCGAGCCCTGGGCATCGCCCCACCCG  
CGCAAGCGCCGCTGGTGGAGCGCGAGAACGCGCGCTGGGCATCGGCGCCCATGATCTTCGGCTTCTGGCGCGCGCGCTCCACCATGG  
GGCCGCTCCATCACCTGACCGTGCAGGCCCGCCAGCTGTTCGGCATCGTGAGCAGCAGTCCAACTGCTGGCGCCATCGAGGCC  
CAGCAGCACTGCTGCAGTGACCGTGTGGGCATCAAGCAGCTGCAGGCCCGCGTGTGGCGTGGAGCGCTACCTGAAGGACCAAGATT  
CCTGGCCCTGTGGGCTGCTCCGGCAAGATCATCTGCACCAACCGCGCTGGAACTCCACCTGGTCCAAACCGCTCCTTCGAGGAGATCT  
GGAACAACATGACCTGGATCGAGTGGAGCGCGAGATCTCCAACATACCAACCAAGATCTACGAGATCTTGACCGAGTCCCAAGAACCAAGCAG  
GACCGCAACGAGAACCTGCTGGAGCTGGACAAGTGGCCCTCCCTGTGGAACCTGGTTCGACATCAACCACTGGCTGTGGTACATCAAGAT  
CTTCATCATGATCGTGGCGGCTGATCGGCCCTGCGATCATCTTCGCCGTGCTGTCCATCGTGAACCGCGTGGCCAGGGCTACTCCCCC  
TGTCCCTCCAGACCCCAACCAACAGCGGAGCCCGACCGCCCGGAGCGCATCGAGGAGGGCGCGGCGGAGCGGCGGCGGCGGCGGCTCC  
GTGGCCCTGGTGTCCGGCTTCCTGGCCCTGGCTGGACGACCTGGCTCCCTGTGCTGCTTCTCCCTACCAACCGCTGCGCGACTTCATCCT  
GATCGCCCGCGCACCGTGGAGCTGCTGGCCACTCCTCCCTGAAGGGCTCGCCCGCGCTGGAGGGCTTGAAGTACCTGGGCAACCTGC  
TGCTGTACTGGGCGCAGGAGCTGAAGATCTCCGCCATCTCCCTGCTGGACGCCACCGCCATCGCCGTGGCGGCTGGACCGACCGCGTGATC  
GAGGTGGCCCAAGGGCGCCCTGGCGCGCCCATCTCTGCACATCCCCCGCCGCTCCGCGAGGGCTGGAGCGCGCCCTGCTGTAA

90/178

**Fig. 54B**

2003 CON 02 AG Env.seq.opt

ATATGCGCGTGATGGCATCCAGAAAGAACTACCCCTGCTGTGGCGCTGGGCATGATCATCTTCTGGATCATGATCATCTGCAACGCCGAGAA  
CCCTGTGGGTGACCGTGTACTACGGCGTGCCTGTGGCGGACGCCGAGACCACCTGTCTGGCCCTCCGACGCCAAGGCTTACGACACCG  
AAGGTGCACAACCGTGTGGGCCACCCACGCTGCGTGCACCCGACCCCAACCCAGGAGATCCACCTGGAGAACGTGACCGAGAACTTCAAC  
ATATGTGAAGAACAACATGGTGGAGCAGATGCACGAGACATCATCTCCCTGTGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCT  
GTGCGTGACCTGGACTGCCACAACAACATCACCAACTCCAACACCAACCAACGCGCGGAGATCAAGAACTGCTCCTTCAACATGA  
CCACCGAGCTGCCGCAAGAAGCAGAAGTGTACGCCCTGTTCACCGCTGGACGTGTGCAGATCAACAAGAACTCCCACTACCGTACCG  
CTGATCAACTGCAACACCTCCGCCATCACCCAGGCTGCCCAAGTGTCTTCGAGCCCATCCCCATCCACTACTGCGCCCGCGCGCTT  
CGCCATCCTGAAGTCAACGACAGGAGTTCACCGGACCGGCCCTTCAAGAACTGTCCAACGTGCAGTGCACCCACGGCATCAAGCCCG  
TGGTGTCCACCCAGCTGCTGAACGGCTCCCTGGCCGAGGAGAGATCGTGATCCGCTCCGAGAACATCACCAACACGCCAAGACCATC  
ATCTGTGCAGCTGTTGAAGCCCTGTGAAGATCAACTGCAACCCGCCCAACAACACCCGAACTCCGTGCGCATCGGCCCGCGCAGACCTT  
CTACGCCACCGCGGACATCATCGCGGACATCCGCCAGGCCACTGCAACGTGTCCGCAACCAAGTGGAAACAACACCTGCGAGCAGTGGCCA  
CCCAAGTGCACAAGTACTTCAACAAGACCATCATCTTCGCCAACCCCTCCGGCGGACCTGGAGATCACCAACCACTCCTTCAACTGCGGC  
TTCCTACTGCAACACCTCCGAGCTGTTCAACTCCACCTGGAACCTCCACCTGGAAACAACACCGAGAAGTGCATCACCCCTGCAGTG  
CCGCAATCAAGCAGATCGTGAACATGTGGCAGAAGTGGGCCAGGCCATGTACGCCCCCCCATCCAGGGCGTGATCCGCTGCGAGTCCAACA  
TTACCGGCCCTGCTGCTGACCCCGACGGCGCAACAACAACTCCACCAACGAGACCTTCGCCCCCGCGGGCGGACATGCGCGACAACCTGG  
TCGCTCCGAGCTGTACAAGTACAAGTGGTGAAGATCGAGCCCTTGGCGTGGCCCCACCCCGCGCAAGCGCCGCTGGTGGAGCGCGAGAA  
GCGCGCGCGCTGGCGCGCTGTTCTGGCTTCTTGGCGCGCGCGGCTCCACCATGGCGCGCCCTCCCATCACCTGACCGGTGCAGG  
CCCCCGCAGCTGTCTCCGGCATCGTGCAGCAGCAGTCCAACCTGTGCGGCCCATCGAGGCCACGACGACCTGCTGAAGCTGACCGTGTGG  
GGCATCAAGCAGCTGCAGGCCCGCGTGTGCCCTGGAGCGCTACCTGAAGGACCGCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCT  
GGATCTGCACCAACCAACCGTGCCCTGGAACCTCCTCTGTTCCAAACAGACCTACAACGACATCTGGACAACATGACCTGSGTGCAGTGGGACA  
AAGGAGATCTCCAACATACACCGACATCATCTACAACCTGATCGAGGAGTCCCAAGAACCGAGAGAACAGCAGGACCTGCTGGCCCTG  
GGACAAGTGGCCCTCCCTGTGGAACCTGTTGACATCAACCACTGGCTGTGTTACATCAAGATCTTCAATCATGATCGTGGCGGCCCTGATCGG  
CCCTGCGCATCGTGTTCGCCGTGCTGACCATCATCAACCGCTGCGCCAGGCTACTCCCCCTGTCTTCCAGACCTTGACCCACCAACGAG  
GCGGAGCCCGAGCCCGAGCGCATCGAGGAGGCGGCGGAGCAGGACCGCATCCGCTCCGTGCGCCCTGGTGTCCGGCTTCTTGGCCCTG  
CCCTGGAGCAGCTGCGCTCCCTGTGCCCTGTTCTCTTACCAACCGCTTGCAGCTTCGTGTGCTGATCGCCCGCCGACCGTGGAGCTGCTGGG  
CCACTCTCTCCCTGAAGGCCCTGCGCCTGGCTGGAGGCCCTGAAGTACCTGGGCAACCTGCTGTCTACTGGGGCCAGGAGCTGAAGAACT  
CCGCGCATCAACCTGTGGACACCATCGCCATCGCCGTGGCCAACTGGACCGACCGGTGATCGAGATCGGCCAGCGCGCGCGCCATC  
TGAACATCCCCCGCGCATCCGCCAGGGCTTGGAGCGCGCCCTGCTGTAA

91/178

Fig. 55A

2003 CON 03 AB Env

MRVKEIRKHLWRWGTLFLGLMLICSATENLWTVVYGVVWKEATTLFCASDAKAYSKEVHNWVATYACVPTDPSQEIPIENVTFENFMG  
 KNNMVEQMHEDEIISLWDQSLKPCVKLTPLCVTLNCTDLKKNVTSTNTSSIKMEMKNCSEFNITDLRDKVKKEYALFYKLDVVQIDNDSYRL  
 ISCNTSVVTOACPKISFEPIPIHYCAPAGFAILKCNCKKFNSTVQCTHGIKPVVSTQLLNGSLAEEVIRSVNFTDNTKTI  
 VOLKEPVEINCTRPNNNTRKGIHIGPGRAFYATGDIIGDIRQAHCNISITKWNNTLKQIVIKLRKQFGNKTIVFNQSSGGDPEIIVMHSFNCG  
 GEFYCNTTKLFNSTWNGTEELNNTGDIIVTLPCRKQIINMWQEVGKAMYPPIAGQIRCSSNITGLLLTRDGGNQSNVTEIFRPGGGDMR  
 DNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGAVFLGELGAAGSTMGAASITLTVOARQLLSGIVQQQNNLLRAIEAQHLLQL  
 TVWGIKQLQARVLAVERYLKDQQLGIWGCSGKLICTTAVPWNNTSWNSKSLDEIWNMTWMEWEREINNNTGLIYNLIEESQNZQEKNEQEI  
 LALDKWASLWNNWFDISKWLWYIKIFIMIVGGLVGLRIIFAVLSIVNRVRQGYSPLSFQTRLPQRPDRPEGIEEGGERDRDTSIRLVNGF  
 LALIWDRLSLCLFIYHHLRDLILLIARIVELLGRRGWEALKYWNLLQYWIQELKSSAINLIDTIAIAVAGWTDVIEIGQRFCAIRNIP  
 RRIRQGAEKALQ\$

Fig. 56A

2003 CON 04 CPX Env

MRVMGIQRNYPHLWEWGTLILGLVIIICSASKNLWTVVYGVVWWRDAETTPFCASDAKAYDKEVHNWATHACVPTDNPQEIALKNVTFENF  
 NMWKNMVEQMHEDEIISLWDEGLKPCVKLTPLCVALNCNATINNSTKTNSTEEIKNCSEFNITTEIRDKKKKEYALFYRLDIVPINDSANN  
 SINSEYMLINCNASTIKQACPKVTFEPIPIHYCAPAGFAILKCNCKKFNSTVQCTHGIKPVVSTQLLNGSLATEGVVIRSKNF  
 TDNTKNIIVQLAKAVKINCTRPNNNTRKSVHIGPGQWYATGEIIGDIRQAHCNISGNDWNETLQKIVEELRKHFPNKTIIFAPSAGGDLEI  
 TTHSFNCGGEFFYCNTSELFNSTYMNSTNTINKTITLPCRKQIIVSMWQEVGQAMYPPIAGSINCSSDITGIILTRDGGNNNTNNETFR  
 PGGDMRDNRSELYKYKVVKIEPVGVAPTRARRRVVQREKRAVGIGAVFLGELGAAGSTMGAASITLTVOARQLLSGIVQQQNNLLRAIEA  
 QQHLLRLTVWGIKQLQARVLALESYLKDQQLGIWGCSGKLICTTNPWNSSWSNKSNDIWDNMTWLQWDKEINNNTQIIYELLEESQNZQ  
 EKNEQDLALDKWANLWNNWFNISNWLWYIKIFIMIVGGLIGIRIIFAVLSIVNRVRQGYSPLSQTLIPTTQRPDRPEGIEEGGERDRSR  
 SIRLVNGFLPLIWDRLNLCFSYRHLRNLILLIIVARTVELLGIWGWEALKYILWNNLLLYWGQELRNSAINLDDTTAIAVAEGTDRIIEAVQRA  
 CRAIRNIPRIRQGLERALL\$

92/178

Fig. 55B

2003 CON 03 AB Env. seq. opt

ATGCGCGTGAAGAGATCCGCAAGCACCTGTGGCGTGGGACCCCTGTCTCTGGGCAATGCTGATGATCTGCTCCGCCACCGAGAACCTGTG  
GGTGACCGTGTACTACGGCGTGGCGGTGCGGTGGAAGGAGGCCACACACCCCTGTCTGGCGCTCCGACGCCAAGGCCCTACTCCAAGGAGGTGC  
ACAACGTGTGGCCACCTACGCTGCGTGGCCACCGACCCCTCCCGCCAGGAGATCCCGCTGGAGAACGTGACCCGAGAACTTCAACATGGGC  
AAGAACAACATGCTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGACCACTCCCTCCATCAAGATGATGGAGATGAAGAACTGCTCTTCAACA  
GACCTGAACATGCACCGACCTGAAGAAGAACGTGACCTCCACCAACACCTCTCTACAAAGTGGACGTGGTGCAGATCGAACACGACTCCTACCGCTG  
TACCAACCGACCTCGCGGACCAAGTGAAGAAGGATACGCTTCTACAAAGTGGACGTGGTGCAGATCGAACACGACTCCTACCGCTG  
ATCTCTGCAACACCTCCGTGGTGAACCCAGGCTTGCCCAAGATCTCCTTCGAGCCCCATCCCCATCCACTACTGCGCCCCCGCGGCTTCGC  
CATCCTGAAGTGAACGACCAAGTTCAACGGCACCGGCCCTGACCAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCCGTGG  
TGTCCACCCAGCTGCTGAACGGCTCCCTGGCCGAGGAGGTGGTATCCGTCCGTGAACCTTCAACGACAAACCAAGACCATCAAGCCCCGTGG  
GTGACGTGAAGGAGCCCGTGGAGATCAACTGCACCCGCCCAACAAACACCCGCAAGGCAATCCACATCGGCCCGCGCGCTTCTA  
CGCAACCGCGGACATCATCGCGGACATCCGCCAGGCCACTGCAACATCTCGATCACCAAGTGAACAAACACCTGAAGCAGATCGTGATCA  
AGCTGGCAAGCAGTTCGGCAACAAGACCATCGTGTTCACCAAGTCTCCGGCGGACCCCGAGATCGTGATGCACTCCTTCAACTGCGGC  
GGCGAGTTCTTCTACTGCAACACCAACCAAGCTGTTCAACTCCACCTGGAACGGCACCGAGAGCTGAACAACACCGAGGGCGACATCGTGAC  
CCTGCCCTGCCGATCAAGCAGATCATCAACATGTGGCAGGAGGTGGCAAGGCCATGTACGCCCCCCCATCGCCGCCAGATCCGCTGCT  
CCTCCAACATCACCGGCTGCTGTGACCCGCGACGGCGCAACCAAGTCCACGTGACCGAGATCTCCGCCCGCGCGGCGGACATGCGC  
GACAACTGGCGCTCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTGGGCGTGGCCCCCACCAGGCCCAAGCCCGCGTGGTGA  
GCGGAGAAGCGCGCGTGGCATCGGCGCGTGTCTCTGGCTTCTGGCGCGCGCGCTCCACCATGGGCGCGCGCTCCATCACCTGA  
CCGTGACAGGCCCGCAGCTGCTGTCCGCGCATCGTGACGACGAGAACAACTGCTGCGCGCCATCGAGGCCACGACACCTGCTGCAGCTG  
ACCGTGTGGGCAATCAAGCAGCTGCAGGCCCGCTGCTGGCGTGGAGCGCTACCTGAAGGACCAAGCTGCTGGGCATCTGGGGCTGCTC  
CGCAAGCTGATCTGCACCAACCGCGTGGCTGGAACACCTCCTGGTCCAACAAGTCCCTGGACGAGATCTGGAACAACATGACCTGGATGG  
AGTGGGAGCGCGAGATCAACAATAACCGGCTGATCTACAACCTGATCGAGGAGTCCAGAACCCAGAGGAGAACGAGCAGGAGATC  
CTGGCCCTGGACAAGTGGCCCTCCCTGTGGAACCTGGTTCGACATCTCCAAGTGGCTGTGTAACATCAAGATCTTCAATCATGATCGTGGCGG  
CCTGGTGGCCCTGCGCATCATCTTCGCCGTGCTGCCATCGTGACCGGCTGGCCAGGGCTACTCCCCCTGTCTTCCAGACCCGCTGC  
CCACCCAGCGCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGAGCGGACCGCGACACTCCCATCCGCTGGTGAACGGCTTC  
CTGGCCCTGATCTGGGACGACCTGGCTCCCTGTGCTGTTCACTACCAACCTGCTGCGGACCTGCTGCTGATCGCGCGCGCGCATCGTGGA  
GCTGCTGGCGCGCGGCTGGGAGGCCCTGAAGTACTGTTGAACCTGCTGCACTGATCGAGGAGTGAAGTCTCCGCGCATCAAC  
TGATCGACACCATCGCCATCGCCGTGGCGGCTGGACCGGCGGTGATCGAGATCGGCCAGCGCTTCTGCCGCGCCATCCGCAACATCCCC  
CGCCGATCCGCCAGGGCGCGGAGAGGCCCTGCAAGTAA



94/178

Fig. 57A

2003 CON 06 CPX Env

MRVKGIQKNWQH<sup>1</sup>HLWKWGTLLILGLV<sup>1</sup>IICSASNMMWTVYYGVPAWEDADTILFCASDAKAYSAEKHNWVWATHACVPTDPNPQOEIALENVTENF  
 NMWKNHMEQMHEDIISLWDESLKPCVKLTPLCVTLNCTNVTKNNTKIMGREEIKNCSEFNVTEIRDKKKKEYALFYRLDVVPIDDDNNNSY  
 RLINCNASTIKQACPVSFEPIPIHYCAPAGFAILKCRDKNFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEII<sup>1</sup>IKSENLT<sup>1</sup>DN<sup>1</sup>T<sup>1</sup>KT  
 IIVQLNKSVEIRCTRPNNNTRKSI<sup>1</sup>SFGPGQAFYATGDIIGDIRQAHCVSRTDWNMLQNV<sup>1</sup>TAKLKELFNKNIT<sup>1</sup>FNSSAGGDLEIT<sup>1</sup>THSFNC  
 GGEFFYCNTSQLFNSTRPNETNTITLPCKIKQIVRMWQVGOAMYAPPIAGNITCTSNITGLLLTRDGNNDSETFRPGGDMRDNRSELY  
 KYKVVKIKPLGIAPTRARRRVGREKRAVGLGAVFLGFLGTAGSTMGAASITLTVQVRQLSGIVQQSNLLRAIEAQ<sup>1</sup>QHLLQLTVWGIKQL  
 QARVLAVERYLKDQQLGIWCGSKLICPTNVPWNASWSNKT<sup>1</sup>YNEIWDNMTWIEWDREINNYTQ<sup>1</sup>QIYSLIEESQ<sup>1</sup>QNEQD<sup>1</sup>LLALDKWAS  
 LMSWEDISNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSLQTLIPNPTGADRPGEIEEGGEGQGRTRSIRLVNGFLALAWDDL  
 RSLCLFSYHRLRDFVLIAARTVETLGHARGWEILKYLGNLVCYWGQELKNSAISLLDTTAIAVANWTDRVIEVVQRVFRAFLNIPRIRQGF  
 RALL\$

Fig. 58A

2003 CON 08 BC Env

MRVRGTRRN<sup>1</sup>YQ<sup>1</sup>Q<sup>1</sup>W<sup>1</sup>IWGVLGFWMLMICNVEGNLW<sup>1</sup>VTVYYGV<sup>1</sup>PVWKEAKTTLFCASDAKAYETE<sup>1</sup>VHNVWATHACVPTDPNPQOEIVMENVTENF  
 NMWNNDMVNQM<sup>1</sup>HEDVISLWDQSLKPCVKLTPLCVTLECTNVSSNGNGTYNETYNESVKEIKNCSEFNATLLRDRKKT<sup>1</sup>VYALFYRLDIVPLND  
 ENSGKNSSEYYRLINCNTSAITQACP<sup>1</sup>KVT<sup>1</sup>FDPIPIHYCTPAGYAILKCN<sup>1</sup>DKKFN<sup>1</sup>GTGQCHNVSTVQCTHGKIPVSTQLLNGSLAEREII  
 RSENLTNNVKTIIIVHLNQSV<sup>1</sup>EIVCTRPNNNTRKSI<sup>1</sup>RIGPGQTFYATGDIIGDIRQAHCNISKDKWYETLQ<sup>1</sup>RVSKKLAEHFPNKT<sup>1</sup>IKFASSSG  
 GDLEITTHSFNCRGEFFYCNTSGLFNGTYMNGTNNSSIIITPCRIKQIINMWQEVGRAMYAPPIEGNITCKSNITGLLLV<sup>1</sup>RDGGRTESNNT  
 EIFRPGGDMRN<sup>1</sup>NRNELYKYKVVEIKPLGVAPTAAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLR  
 AIEAQ<sup>1</sup>QHMLQLTVWGIKQLQTRVLAIERYLK<sup>1</sup>DQQLGIWCGSKLIC<sup>1</sup>TAVPWNSSWSNKSQ<sup>1</sup>QEIWDNMTW<sup>1</sup>MQWDKEISNYTNTIYRLLEDS  
 QNQQERNEKDLLALDSWKNLMSWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQILTPNPGGPRGLGRIEEEGEGEQD  
 KTRSIRLVNGFLALAWDDLRLNLC<sup>1</sup>FSYHRLRDFILLTARGVELLGRNSLRGLQRGWEALKYLGSLVQYWGLELKKSTISLVD<sup>1</sup>TIAIAVAEGT  
 DRIINIVQGICRAIHNI<sup>1</sup>PRIRQGF<sup>1</sup>EALQ\$



95/178

Fig. 57B

2003 CON 06 CPX Env. seq. opt  
 ,ATGCGCGTGAAGGGCATCCAGAAAGAACTGGCAGCACCTGTGGAAGTGGGGCACCCCTGATCCTGGGCGCTGGTGATCATCTGCTCGGCTTCCAA  
 CAACATGTGGGTGACCGTGTACTACGGCGTGTCCCGCTGGAGGACCGCGACACCATCCTGTTCTGCGCTCCGACGCCAAGGCTACTCCG  
 CCGAGAAGCAACAGTGTGGCCACCCACCGCTGCGTGGCCACCGACCCCAACCCCGAGGATCGCCCTGGAGAACGTGACCGAGAACTTC  
 AACATGTGGAAGAACACATGGTGGAGCAGATGCACGAGACATCATCTCCCTGTGGACGAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCCCTGAACCTGCACCAACCTGACCAAGAACAAACACCAAGATCATGGGCGCGAGGAGATCAAGAACTGCTCCTTCAACG  
 TGACCACCGAGATCCGCGACAAGAGAGAGTAGCCCTGTTTACCGCTGGAGTGGTGCCATCGACGACAACAACTCCTAC  
 CGCTGATCAACTGCAACGCTCCACCATCAAGCAGGCTGCCCCAAGGTGTCTTCGAGCCCATCCCATCCACTACTGCGCCCCCGCGG  
 CTTGCGCATCCTGAAGTGGCGGACAAAGAACTTCAACGGCACCGGCCCTGTCAAGAACGTGTCCACCGTGCAGTGCACCCCGGCTCAAGC  
 CCGTGGTTCACCCAGCTGCTGTAAACGGCTCCCTGGCCGAGGAGAGATCATCAAGTCCGAGAACCTGACCGCAACACCAAGACC  
 ATCATCGTGCAGCTGAACAAGTCCGTGGAGATCCGCTGCACCCGCCCAACAAACAGTGTCCCGACCGACTGGAACAACATGCTGCAAGACGTGA  
 CTTTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGTCCCGACCGGACCTGGAGATCAACACCATCACCCTGCTCAACTGC  
 CCGCAAGCTGAAGGAGCTGTTCAACAAGAACATCACTTCACTCCCGCGCGGACCTGGAGATCAACACCATCACCCTGCTCAAGATCAA  
 GCGGCGAGTTCTTACTGCAACACCTCCAGCTGTTCAACTCAACCGCCCCCAACGAGACCAACACCATCACCCTGCAACATCAACGATCAA  
 GCAGATCGTGCAGATGTGGCAGCGGTGGCCAGGCCATGTACGCCCCCCCATCGCCGGCAACATCACTGCACCTCCAACATCAACGATCAA  
 TGCTGTGACCCGCGACGCAACAACAGCTCCGAGACCTTCCGCCCGCGGCGCGACATGCGCGACAACTGGCGTCCGAGCTGTAC  
 AAGTACAAGGTGGTGAAGATCAAGCCCTGGGATCGCCCCACCGCGCGCGCGCGCTGGTGGCGCGGAGAACCTGCGCTCCGAGCTGTAC  
 GCGCGCGTGTCTGGCTTCTTGGCACCGCGCTCCACCATGGGCGCGCGCTCCATCACTGACCGTGCAGGTGCGCGAGCTGTGT  
 CCGGCATCGTGCAGCAGCTCCAACTGTGTGCGGCCATCGAGGCCAGCAGACCTGTGACGTGACCGTGTGGGATCAAGCAGCTG  
 CAGGCCCGCTGCTGGCGGTGAGCGCTACCTGAAGACCAAGCAGCTGTGGGATCTGGGCTGTCCGGCAAGCTGATCTGCCCCACCAA  
 CGTGCCCTGGAACGCCCTCTGGTCCAACAAGACCTACAACGAGATCTGGGACAACTGACCTGGATCGAGTGGGACCGCGAGATCAACAAT  
 ACACCCAGCAGATCTACTCCCTGATCGAGAGTCCAGAACGAGAGAGAACGAGCAGGACCTGTGGCCCTGGACAAGTGGGCTCC  
 CTGTGCTCCTGGTTCGACATCTCAACTGGTGTGGTACATCAAGATCTTCAATGATCGTGGGCGCTGTATCGGCTTGGCATCTGTGT  
 CGCCGTGCTGTCCATCGTGAAACCGGTGCGCCAGGGCTACTCCCCCTGTCCCTGCAGACCTGATCCCCAACCCCAACCGCGCGACCGCC  
 CCGCGAGATCGAGGAGGCGGCGGAGAGGGCGGACCCGCTCCATCCGCTGGTGAACGGCTTCTTGGCCCTGGCTGGGACGACCTG  
 CGCTCCCTGTGCTGTTCTCTACACCGCTGCGGACTTCGTGTGATCGCCCGCGCACCGTGGAGACCTTGGGCGCACCGCGGCTGGGA  
 GATCTGAAGTACCTGGGCAACCTGGTGTGCTACTGGGCGCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAACCGGCTGCGG  
 TGGCCAACTGGACCGACCGGTGATCGAGGTGGTGCAGCGCTGTCCGCGCTTCTGTAACATCCCCCGCGCGCATCCGCCAGGGCTTCGAG  
 CGCGCCCTGCTGTAA



96/178

Fig. 58B

2003 CON 08 BC Env seq. opt  
 ATGCGGTGGCGGCAACCGCGCAATACACAGAGTGGTGATCTGGGGCGTGCTGGGCTTCTGGATGCTGATGATCTGCAACGCTGGAGGG  
 CAACCTGTGGGTGACCGGTGTAACGGCGTGGCGGCGTGGAGAGGCGCAAGACCAACCTGTTCTGCGCTCCGACGCCAAGGCTACGAGA  
 CCGAGGTGCACAAAGTGTGGGCCACACCGCTGGTGGCCACCGACCCCAACCCAGGAGATCGTATGGAGAACGTCGACCCGAGAACTTC  
 AACATGTGGAACAACGACATGTGTGAACCAAGATGCACGAGACGTGATCTCCCTGTGGGACCAATCCCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGACCCCTGGAGTGACCAACGTTCTCCAAACGGCAACGGCACCTACAAACGAGACCTACAAACGAGTCCGTGAAGAGATCAAGA  
 ACTGCTCCTTCAACGCCACCAACCTGTGCGGACCGCAAGAACCGGTGTACGCCCTGTTCTACGCCCTGGACATCCGTCCTTGAACGAC  
 GAGAACTCCGGCAAGAACTCCTCCGAGTACTACCGCTGATCAACTGCAACACCTCCGCCATCAACAGGCTGCCCAAGGTGACCTTCCA  
 CCCCATCCCCATCCACTACTGCAACCCCGCGGCTACGCCATCCTGAAGTGCAACGACAAGAGTTCAACGGCACCGGCCAGTGGCCACAACG  
 TGTCCACCGTGACGACCCACCGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGTGAACGGCTCCCTGGCCGAGCGCGGAGATCATCATC  
 CGCTCCGAGAACCTGACCAACAAACGTGAAGACCATCATCGTGACCTGAACCAAGTCCGTGGAGATCGTGTGACCCCGCCCAACAAACAC  
 CCGCAAGTCCATCCGATCCGATCGGCCCGCGGAGACCTTCTACGCCACCGCGGACATCATCGCGGACATCCGCCAGGCCACTGCAACATCTCCA  
 AGACAAAGTGTAGAGACCCCTGACGCGCGTGTCCAAGAGCTGGCCGAGCACTTCCCAACAAAGACCATCAAGTTCGCTCCTCCTCCGCG  
 GGCACTGTGGAGATCACCAACCTCTTCAACTGCCCGCGGAGTTCTTCTACTGCAACACCTCCGGCTGTTCACGGCACCTACATGAA  
 CGGACCAACAATCCTCCTCCATCATCAACCTCCCTGCCGATCAAGCAGATCATCAACATGTGGCAGGAGTGGCCGCGCCATGTACG  
 CCCCCCATCGAGGGCAACATCACTGCAAGTCCAACATCAACCGCTGCTGTGGTGCGGACGGCGGCCGACCGAGTCCAACAACACAC  
 GAGATCTTCCGCCCGCGGCGGACATCGCGCAACAACCTGGCGCAACGAGCTGTACAAGTACAAGTGTGGAGATCAAGCCCTTGGGCGT  
 GGCCCCACCGCGCGCAAGCGCGCGTGGTGGAGCGCGGAGAGCGCGCGTGGCCCTGGCCCTGGCCCTGTTCTTCTGGCTTCTTGGCGCGCGG  
 GCTCCACCATGGCGCGCTCCATCACTGACCGTGCAGGCGCGCGCATCAAGCAGTGCAGACCCCGTGGCCCTGGAATCCTCTGTGTCACAAAGTCCC  
 GCCATCGAGGCGCAGCAGCATGTGCAAGTGCAGTGCAGTGGGCGATCTCGCAACCGCGCTGGCCCTGGAATCCTCTGTGTCACAAAGTCCC  
 GGACCAAGCATCTGGGCAACATGACCTGGATGCAGTGGGCAAGGAGATCTCCAACTACACCAACACCATCTACCCCTGCTGGAGGACTCC  
 AGCAGGAGATCTGGGCAACATGACCTGGATGCAGTGGGCAAGGAGATCTCCAACTACACCAACACCATCTACCCCTGCTGGAGGACTCC  
 CAGAACCAAGAGCGCAACGAGAGGACCTGCTGGCCCTGGACTCCTGGAAAGAACTGTGGTCTGTGTTCTGACATCAACAACTGGCTGTG  
 GTACATCAAGATCTTCAATCATGATCGTGGCGGCTGATCGGCTGGGCGCATCATCTTCCGCGTGTGTCCATCGTGAACCGCGTGGCGCAGG  
 GCTACTCCCCCTGTCTTCCAGATCCTGACCCCAACCCCGCGCGCCCGCTGGCCGCGCATCGAGGAGGAGGGCGGCGGAGCAGGAC  
 AAGACCCGCTCCATCCGCTGGTGAACGGCTTCTGGCCCTGGGACGACCTGCGCAACCTGTGCTGTGTTCTCTACCAACCGCTGCG  
 CGACTTCATCTGTCACCCCGCGGCTGGAGTGTGGCGCGCAACTCCCTGCGCGGCTGACGCGCGGCTGGAGGCGGCTGAAGTACC  
 TGGGCTCCCTGGTGCAGTACTGGGGCTTGAAGTCCACCATCTCCCTGTGGACACCATCGCCATCGCCGTGGCGGAGGCGAC  
 GACCGCATCATCAACATCGTGCAGGGCATCTGCCCGCGCCATCCACAAACATCCCCCGCGCATCCGCCAGGGCTTCGAGGCGCGCTGCAGTA

A

97/178

Fig. 59A

2003 CON 10 CD Env

MRVMGIQRNCQQWIIWILGFWMMLICNATGNLWVTVYGVVWKETTTLFCASDAKAYKAEAHNIWATHACVPTDPNPQEIPLENVTFENF  
 NMWKNMGVDMQMHEDIISLWDQGLKPCVKLTPLCVTLNCSVDNATNSATNTVAGMKNCSEFNITTEIRDKKKQEYALFYKLDVVQIDGSNTSY  
 RLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNDDKKFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEIIIRSENLTDNAKT  
 IIVQLNESVTINCTRPNNNTRKSIRIGPGQTFYATGDIIGNIRQAYCNISGTEWNKTLOQVAKKGLDLNKTIIIFKPSSGGDPEITHTTFN  
 CGGEFFCYNTSKLFNSWTSNNTGTSTITLPCRIKQIINMWQGVGKAIYAPPIAGLINCSSNITGLLTRDGGANNSETFRPGGDMRDNW  
 RSELYKYKVVKIEPLGLAPTKAKRRVVEREKRAIGLGAFLGFLGAGSTMGAASLTITVQARQLLSGIVQQQNNLLRAIEAQHLLQLTVW  
 GIKQLQARVLAVESYLKDQQLGIWGCSGKHICTTNVPWNSSWSNKSLEEIWDNMTWMEWEREIDNYTGLIYSLIEESONQOEKNEQELLQL  
 DKWASLWNWFSITNLWYIKIFIMIVGGLIGLRIVFAVLSLVNRVRQGYSPLSFQTLPPAPRGDPREGIEEGEGGQGRSIRLVNGFSAL  
 IWDDLRLNCLFSYHRLRLDILLIATRIVELLGRRGWEAIKYLWNLQYWIQELKNSAISLLDTTATAIAVAEGTDRAIEIVQRAVAVLNIPTRI  
 RQGLERALLS

Fig. 60A

2003 CON 11 CPX Env

MRVKETQRNWHNLWRWGLMIFGMLMICNATENLWVTVYGVVWKDADTTTLFCASDAKAYSTEKHNWATHACVPTDPNPQEIPLENVTFENF  
 NMWKNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVKNATNTTVEAAEIKNCSEFNITTEIKDKKKQEYALFYKLDVVPINDNNNSIY  
 RLINCNVSTVKQACPKVTFEPIPIHYCAPAGFAILKCNDDKKFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEGEVRISENFTNNAKT  
 IIVQLNSSVRINCTRPNNNTRKSIHIGPGQAFYATGDIIGDIRQAHNCISRAEWNNTLOQVAKQLRENFNKTIIIFNPPSGGDLITTHSFNC  
 GGEFFCYNTSRLFNSTWNNDRNDTKQMHITLPCRIKQIVNMWQVRVGQAMYAPPIQGIKIRCNSTNITGLLTRDGGNNNTNETFRPTGGDMRD  
 NWRSELYKYKVEIKPLGVAPTRAKRRVVEREKRAVGIGAVLLGFLGAGSTMGAASITLTVQARQLLSGIVQQQNNLLKAIEAQHLLKLT  
 VMGIKQLQARVLAVERYLKDQQLGIWGCSGKLICTTNVPWNFSWSNKSDEIWDNMTWIEREINNYTQTIYTLLEESQNOQEKNEQDLL  
 ALDKWASLWNWFDISNLWYIKIFIMIVGGLIGLRIIFAVLSIVNRCRQGYSPLSFQTLTPNHKEADRPGGIEEGGEGQDRTRSIRLVSGFL  
 ALAWDDLRNLCFSYHRLRDFILIAARIVETLGRRGWEILKYLGNLAQYWGQELKNSAISLLNATAIAVAEGTDRIIEVVHVRVLRILHIPR  
 RIRQGFERALLS

98/178

Fig. 59B

2003 CON 10 CD Env. seq. opt

ATGCGGTGATGGGCAATCCAGCGCAACTGCCAGTGGTGGATCTGGGGCATCTGGGCTTCTGGATGCTGATGATCTGCAACGCCACCCGG  
 CAACCTGTGGTGACCGGTGTAACGGCGTGGGAGGAGACCAACACCCCTGTTCTGCGCTTCTGCGCCCTCCGACGCCAAGGCTTACAAGG  
 CCGAGGCCACAACATCTGGGCCACCCACCGCTGGTGGCCACCGACCCCAACCCAGGAGATCGTGTGGAGAACGTGACCGAGAATTCTC  
 AACATGTGGAAGAACGGCATGGTGGACAGATGCACGAGACATCATCTCCCTGTGGGACCCAGGGCTGAAGCCCTGCGTGAAGCTGACCCC  
 CCTGTGCGTGAACCTGAACTGCTCCGACGTGAACGCCAACCAACTCCGCCACCAACCCGTGGTGGCCGAGATGAGCGGCTCCAACACTCTTCAACA  
 TCACCAAGAGATCCGCGACAAGAGACGAGTACGCCCTGTTCTACAAGTGGACGTGGTGGAGATGAGCGGCTCCAACACTCTTCAACTCTTCA  
 CGCTGATCAACTGCAACACTCCGCCATCACCCAGGCTGCCCAAGGTGACCTTCGAGCCCATCCCATCCACTACTGCGCCCTCCGCCGG  
 CTTCCGCATCCTGAAGTGCAACGACAAGAGTTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGCAACCGCATCAAGC  
 CCGTGGTGTCCACCCAGCTGCTGAACGGCTCCCTGGCCGAGGAGAGATCATCATCCGCTCCGAGAACCTGACCCGACAACGCCAAGACC  
 ATCATCGTGCAGTGAACGAGTCCGTGACCATCAACTGCAACCCGCCCTACTGCAACATCTCCGGCACCGAGTGGAGAACAGACCTGCAGCAGGTGG  
 CTTCTACGCCACCGCGACATCATCGGCAACATCCGCCAGGCTACTGCAACATCTCCGGCACCGAGTGGAGAACAGACCTGCAGCAGGTGG  
 CCAAGAAGCTGGCGGACCTGCTGAACAAGACCAACCATCATCTTCAAGCCCTCTCCGGCGGACCCCGAGATCAACCCACACCTTCAAC  
 TGCGGGGGGAGTTCTTCTACTGCAACACTCCAAGCTGTTCAACTCCTCTGGACCTCCAACACACCGGCAACACCTCCACCATCACCTT  
 GCCCTGCCGATCAAGCAGATCATCAACATGTGGCAGGCGTGGCAAGGCCATCTACGCCCTCCCATCGCGGCTGATCAACTGCTCCT  
 CCAACATCACCGGCTGCTGTGACCCGCGACGGCGGCCCAACAACCTCCGAGACCTTCCGCCCTCCGGCGGCGGACATGCGCGACAACCTGG  
 CGTCCGAGCTGTACAAGTACAAGTGAAGATCGAGCCCTGGGCTGGCCCAAGGCCAAGCCCGCTGCTGGAGCGCGAGAA  
 GCGGCCATCGGCTGGGCGGCTGTTCTGCGGCTTCTGGGCGCGCGGCTCCACCATGGGCGCGGCTTCCCTGACCTGACCGTGCAGG  
 CCGCCAGCTGCTGCCGATCGTGCAGCAGCAACACTGCTGCCGCCATCGAGGCCAGCAGCACCCTGCTGAGCTGACCGTGTGG  
 GGATCAAGCAGCTGCAAGCCCGCTGGCTGGCTGGAGTCTTACTGAAGGACCAAGCAGCTGCTGGGCTGCTCCGGCAAGCA  
 CATCTGCACCAACGTCGCTGGAACCTCTCTGGTCCCAAGTCCCTGGAGGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGC  
 GCGAGATCGACAACCTACACCGGCTGATCTACTCCCTGATCGAGGAGTCCAGAACCAAGCAGGAGAGAACGAGCAGGAGTGTGCAAGCTG  
 GACAAGTGGGCTCCCTGTGGAACCTGTTCTCCATCAACCACTGGCTGTGGTACATCAAGATCTTCAATCATGATCGTGGCGGCTGATCGG  
 CCTGCGCATCGTGTTCGCCGTGCTGCCGTGGTGAACCGCGTGGCCAGGGCTACTCCCTGCTCTCCAGACCTGCTGCCCGGCTTCTCCGCGCTG  
 GCGGCCCGACCGCCCGAGGGCATCGAGGAGGAGGGCGGAGAGGGCCGCGCGGCTCCATCCGCTGGTGAACGGCTTCTCCGCGCTG  
 ATCTGGGACGACCTGCGCAACCTGTGCTGTTCTCTACCAACCGCTGCGGACCTGATCTGATCGCCACCGCATCGTGGAGCTGCTGGG  
 CCGCGCGGCTGGGAGGCCATCAAGTACCTGTGGAACCTGTCGAGTACTGATCCAGGAGTGAAGAACTCCGCCATCTCCCTGCTGGACA  
 CCACCGCCATCGCGCTGGCCGAGGGACCGGACCGGCCATCGAGATCGTGCAGCGGCGCTGCGGCGCTGCTGAACATCCCCACCGCATC  
 CGCCAGGCTGAGCGGCGCTGCTGTA

99/178

Fig. 60B

2003 CON 11 CPX Env. seq. opt

ATGCGCGTGAAGAGACCCAGCGCAACCTGTGGCGCTGGGCGCTGATGATCTTCGGCATGCTGATGATCTGCAACGCCACCGA  
GAACCTGTGGGTGACCGTGTACTACGGCGTCCCGTGTGAAGAGCGCGACACCACTCTTCTGCGCTCCGACGCAAGGCTTACTCCA  
CCGAGAAGCAACAGTGTGGGCCACCCACGCTGCGTGCACCGACCCCAACCCAGAGATCCCTCGAGAACGTGACCGAGAACTTC  
AACATGTGAAGAACAACATGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACGAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
CCTGTGCGTGAACCTGAACCTGACCCGACGTGAAGAACGCCACCAACACCCGCGGAGGCCGCGAGATCAAGAAGCTGCTCCTTCAACATCA  
CCACCGAGATCAAGGACAAAGAAGAGGATACGCCCTGTCTACAAGCTGGAGCTGCTCGCCATCAACGACAAACAACCTCCATCTAC  
CGCTGATCAACTGCAACGTGTCCACCGTGAAGCAGGCTGCCCAAGTGACCTTCGAGCCCATCCCATCTCACTACTGCGCCCCCGCGG  
CTTCGCCATCCTGAAGTGAACGACAAAGAGTCAACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGACCCACGGCATCAAGC  
CCGTGGTGTCCACCCAGCTGCTGTGAACGGTCCCTGGCCGAGGGCGAGGTGCGCATCCGCTCCGAGAACTTCAACAAACGCCCAAGAC  
ATCATCGTGACGTGAACCTCTCCGTGCGCATCAACTGCACCCGCCCAACAAACAACACCGCAAGTCCATCCACATCGGCCCGGACAGC  
CTTACGCCACCGCGACATCATCGGCGACATCCGCCAGGCCACTGCAACATCTCCCGCGCCGAGTGAACACACCTTGCAGCAGGTG  
CCAAGCAGCTGCGGAGAACTTCAACAACCCCTCCCGCGCGGACCTGGAGATCACACCCACTCCTTCAACTGC  
GGCGGAGTTCTTCTACTGCAACACTCCCGCTGTCAACTCCACTGGAAACAACGACACCGCAACGACACCAAGCAGATGCACATCAC  
CCTGCCCTGCCGATCAAGCAGATCGTGAACATGTGGCAGCGGTGGCCAGGCCATGTACGCCCCCATCCAGGGCAAGATCCGCTGCA  
ACTCCAACATCACCGGCTGCTGTGACCCCGACGGCGCAACAACAACCAACGAGACCTTCGCCCCACCGCGCGGACATGCGCGAC  
AACTGGCGCTCCGAGCTGTACAAGTACAAGTGTGGAGATCAAGCCCTGGGCGTGGCCCAACCGCGGCAAGCGCGGTGGAGCG  
CGAGAAGCGCGCGTGGCATCGGCGCGTGTCTGGCTTCTGGCGCGCGCGGCTCCACCATGGCGCGCGCTCCATCACCTGACCG  
TGCAGGCCCGCGAGCTGTGTCCGGCATCGTGCAGCAGTCCAACTGTGAAGGCCATCGAGGCCAGCAGCTGTGGGCTGTGGGCTGCTCCGG  
GTGTGGGCGATCAAGCAGTGCAGGCCGCTGTGGCGGTGAGCGCTACCTGAAGGACCAAGCAGCTGTGGGCTGTGGGCTGCTCCGG  
CAAGCTGATCTGCACCAACAGTGGCTGGAACCTCTCTGGTCCAAAGTCTACGACGAGATCTGGGACAAACATGACCTGGATCGAGT  
GGAGCGCGAGATCAACAACACACCGACCATCTACACCTGTGGAGGAGTCCAGAACCAAGCAGGAGAAAGAACGAGACCTGCTG  
GCCCTGGACAAGTGGCTTCCCTGTGGAACCTGTTCGACATCTCAACTGGCTGTGGAGGAGTCCAGAACCAAGCAGGAGAAAGAACGAGACCTGCTG  
GATCGGCTTGGCATCATCTTGGCGGTGCTGTCCATCGTGAACCGTGCAGGCTACTCCCCCTGTCTTCCAGACCTGACCCCCA  
ACCACAAGGAGCGCGACCGCCCGCGGCGATCGAGGAGGCGCGGCGAGCAGGACCGCACCGCTCCATCCGCTGTGGTCCGGCTTCTG  
GCCCTGGCTGGGACGACCTGCGCAACCTGTGCTTCTCTACCAACCGCTGGCGACTTCATCCTGATCGCGCGCGCATCGTGGAGAC  
CCTGGGCGCGCGGTGGGAGATCCTGAAGTACCTGGGCAACCTGGGCGAGTACTGGGCGCAGGAGTGAAGAACTCCGCGCATCTCCCTGC  
TGAACGCCACCGCATCGCGGTGGCGGAGGACCGGACCGCATCATCGAGGTGTGACCGCGTGTGGCGCGCATCTGACACATCCCCCGC  
CGCATCCGCGCAGGCTTCGAGCGCGCTGCTGTAA

100/178

Fig. 61A

2003 CON 12 BF Env

MRVRGMQRNWOHLGKWLFLGLIICNATENLWVTYYGVVPVWKEATTLFCASDAKSYEREVHNWATHACVPTDPNPQEVLDENVTF  
 DMWKNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCTDANATANATKEHPEGRAIAIQCSEFNTTEVRDKQMKVQALFYRLDIVPISDN  
 NSNEYRLINCNSTSTITQACPKVSWDPIPIHYCAPAGYAILKCNKKFNGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEEEEIIIRSONIS  
 DNAKTIIVHLNESVQINCTRPNNNTRKSIHIGPGRAFYATGDIIGDIRKAHCNVSGTQWNKTLEQVKKKLSYFNTTIKFNSSSGGDPETIM  
 HSFNCRGEFFYCNTSKLFNDTVSNDTIIILPCRIKQIVNMWQEVGRAMYAAPIAGNITCTSNTITGLLLTRDGGHNETNKTETFRPGGGMKDN  
 WRSLEYKYKVVEIEPLGVAPTRAKRQVVKREKRAVGIGALFLGELGAAGSTMGAASITLTVOARQLLSGIVQQSNLLRAIEAQHLLQLTV  
 WGIKQLQARVLAVERYLKDQQLGLWGCCKLICCTNVPNWSSWSNKSQEEIWNMTWMEWEKEINNYSEIYRLIEESQNEKNEQELLA  
 LDKWASLWNWFDISNWLWYIRIFIMIVGGLIGLRIVFAVLSIVNRVRKGYSPLSLQTHIPSPREPDRPEGIEEGGEGQKDRSVRLVNGFLA  
 LIWDDLRSLCLFSYHRLRDLIIIVTRIVELLGRRGWEVLKYWWNLLQYWSQELKNSAISLLNTTAIVVAEGTDRVIEALQRVGRAILNIPRR  
 IRQGLERALL\$

Fig. 62A

2003 CON 14 BG Env

MKAKGTQRNWOHLGKWLFLGLIICNATENLWVTYYGVVPVWKEATTLFCASDAKAYDAEVHNWATHACVPTDPNPQEVLDENVTF  
 NMWENNVDQMQEDIIISLWDQSLKPCVELTPLCVTLNCTDFNNTTNNTTNTRNDGEGEIKNCSFNITTSLRDKIKKEYALFYRLDIVVQMDND  
 NSSYRLTSCNTSIIITQACPKVSFTPIPIHYCAPAGFVILKCNKTFNGTGPCNTVSTVQCTHGIRPVVSTQLLNGSLAEEEEIVIRSKNFTD  
 NAKTIIIVQLKDPFIEINCTRPNNNTRKRITMGPGRVLYTTGQIIIGDIRKAHCNISKTWNNTLGQIVKKLREQFMNKTIVFORSSGGDPETIM  
 HSFNCRGEFFYCNTTQLFNSTWRSNSTWNTDITLPCRIKQIVNMWQKVGKAMYAPPISQIRCSSNITGLLLIRDGGSNNTET  
 RPPGGNMKDNWRSLEYKYKVVKIEPLGVAPTRAKRRVQREKRAVGIGALLFGELGAAGSTMGAASMTLTVOARQLLSGIVQQNNLLRAIE  
 AQOHLQLTVWGIKQLQARVLAVERYLKDQQLGIWGCCKLICCTTVPNWASWSNKSLLDDIWNMTWMEWEEREIDNYTGLIYTLIEQSONQ  
 QERNEQELLELDKWASLWNWFNITNWLWYIKIFIMIIGGLIGLRIVFAVLSIINRVRKGYSPLSFQTLTHHQREPDPRGRIEEGEGEQKDR  
 SIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGLYLWNLNLLYWGRELKNSAINLLDITVAIAVANWTDRA  
 IEVVQRVGRAVLNIPVRIROGLERALL\$

101/178

Fig. 61B

2003 CON 12 BF Env. seq. opt

ATGCGCGTGGGGCATGCAGCGCAACTGGCAGCACCTGGCAAGTGGGGCTGCTGTTCTTGGGCATCCTGATCATCTGCAACGCCACCGA  
GAACCTGTGGGTGACCGGTGTAACGGCGTGCCCGTGTGAAGGAGGCCACCAACACCTGTTCTGGCCCTCCGACGCCAAGTCTCTACGAGC  
GCGAGTGCACAACGTGTGGCCACCCACCGCTGCGTGCCTGAGGAGGTGACCTGGAGAACGTGACCGAGAACTTC  
GACATGTGAAGAACAACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGACCACTCCCTGAAGCCCTGCGTGAAGCTGACCCC  
CCTGTGCGTGACCCCTGAACCTGCACCGACGCCAACCGCCACCGCAAGAGTGAAGTGCAGGCCCTGTTCTACCGCTGGACATCTGTGCCCATCTCCGACAAC  
GCTCCCTTCAACATGACCAACCGAGGTGCGGACAGCAGATGAAGTGCAGGCCCTGTTCTACCGCTGGACATCTGTGCCCATCTCCGACAAC  
AACTCCAACGAGTACCGCTGATCAACTGCAACACCTCCACCATCACCCAGGCTGCCCAAGGTGCTCTGGGACCCCATCCCATCCACTA  
CTGGCCCCCGCGGTACGCCATCCTGAAGTGAACGACAAAGATTCACGGCACCGGCCCTGCAAGAACGTGTCCACCGTGCAGTGCAGTGA  
CCACGGCATCAAGCCCGTGTCCACCCAGTCTGTAACGGCTCCCTGGCCGAGGAGATCATCATCCGCTCCAGAACATCTCC  
GACAAACGCCAAGACCATCATCTGCACTGAACGAGTCCGTGCAGATCAACTGACCCCGCCCAACAACAACACCCGCAAGTCCATCCACAT  
CGCCCCGGCGGCTTCTACGCCAACCGGACATCATCGCGACATCCGCAAGGCCCACTGCAACGTGTCCGGCACCCAGTGGAAACAAGA  
CCCTGGAGCAGGTGAAGAAGCTGCGCTCCTACTTCAACACCATCAAGTTCAACTCCTCTCCGGCGGACCCCGAGATCACCATG  
CACTCCTTCAACTGCCCGCGGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACGACACCGTGTCCAACGACACCATCATCTCCGCTG  
CCGCATCAAGCAGATCGTGAACATGTGGCAGGAGTGGCGCGGCCATGTACGCCGCCCATCGCCGGCAACATCACCTGCACCTCCCAACA  
TCACCGGCTGCTGTGACCCCGGACGGCGGCCAACACGAGACCAACAAGACCGAGACCTTCGCCCGCGGCGGCAACATGAAGGACAAC  
TGGCGCTCCGAGCTGTACAAGTACAAGTGTGGAGATCGAGCCCTGGCGTGGCCGCCAACCCGCGCAAGCGCCAGGTGGTGAAGCGGA  
GAAGCGCGCGTGGCATCGCGGCTTCTTGGGCTTCTTGGGCGCGCGGCTCCACCATGGCGCGCGCTCCATCACCTGACCGTGC  
AGGCCCGCAGCTGTCCGGCATCTGTGCAGCAGCAGTCCAACCTGTGCGGCCATCGAGGCCCAGCAGCCTGCTGCAGCTGACCGTG  
TGGGCATCAAGCAGTGCAGGCCCGCTGCTGGCGTGGAGCGTACCTGAAGCACAGCAGTGTGGGCTGTGGGCTGTCCGGCAA  
GCTGATCTGCACCAACCTGCCCTGGAACTCCTCTGTCCAACAGTCCAGGAGGAGATCTGGAGAACATGACCTGGATGGAGTGGG  
AGAAGGAGATCAACAACCTACTCCAACGAGATCTACCGCTGATCGAGGAGTCCAGAACCCAGCAGGAGAAAGACGAGAGCTGCTGGCC  
CTGGACAAGTGGGCTCCCTGTGGAACTGGTTCGACATCTCAACTGGCTGTGGCAAGGCTACTCCCCCTGTCCCTGCAGACCCACATCCCCTCCC  
CGCCTGCGCATCGTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGGCAGGGCGGCGGAGCAGGGCAAGGACCGCTCCGTGGCTTCCCTGGCC  
CTGATCTGGGACGACCTGGCTCCCTGTGCTGTCTCTACCAACCGCTGCGGACCTGCTGCTGATCGTGACCCCGCATCGTGGAGTGT  
GGCGCGCGGCTGGAGTGTGAAGTACTGGTGAACCTGCTGCAGTACTGGTCCAGGAGTGAAGAACTCCGCCATCTCCCTGTGTA  
ACACCAACCGCATCTGTGTGGCGGAGGGCACCGACCGCGTGTGAGGCCCTGCAGCGCGTGGCGCGGCGCATCTCTGAACATCCCCCGCGC  
ATCCGCCAGGGCTGGAGCGCGGCTGCTGTAA

102/178

Fig. 62B

2003 CON 14 BG Env. seq. opt

ATGAAGGCCAAGGCCACCCAGCGCAACTGGCAGTCCCTGTGGAAGTGGGGCACCCCTGATCCTGGGCCCTGGTGATCATCTGCTCGGCCCTCCAA  
CGACCTGTGGGTGACCGGTGTAACGGCGTCCCGTGTGGAAGAGGCCACCAACCCCTGTCTGGCCCTCCGACGCCAAGGCCCTACGACG  
CCGAGGTGCACAACGTGTGGGCCACCCACGCCCTGGTGGCCACCGACCCCAACCCCAAGAGGTGGCCCTGGAGAACGTGACCCGAGAACTTC  
AACATGTGGGAGAACAAACATGGTGGACCAAGATGCAGGAGGACATCATCTCCCTGTGGACCAAGTCCCTGAAGCCCTGCGTGGAGCTGACCCC  
CCTGTGCGTGAACCTGAACCTGCACCGACTTCAACAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA  
GCTCCTTCAACAATCACCACTCCCTGGCGACAAGATCAAGAAAGGAGTACGCCCTGTCTACAACCTGGACGTGGTGAGATGGACAACGAC  
AACTCCTCCTACCGCCTGACCTCCTGCAACACCTCCATCATCACCCAGGCCCTGCCCAAGGTGCTTCAACCCCATCCCATCTACCTACTG  
CGCCCCCGCGGCTTCGTGATCCTGAAGTGCAACAACAGACCTTCAACGGCACCGGCCCTGCACCAACGTGTCCACCGTGCAGTGCAACCC  
ACGGCATCCGCCCGCGTGTCCACCCAGCTGCTGAACGGCTCCCTGGCCGAGGAGAGATCGTGATCCGCTCCCAAGAACTTCAACCGAC  
AACGCCAAGACCATCATCTGTCAGCTGAAGGACCCCATCGAGATCAACTGCACCCCGCCCAACAACAACCAACCAACCAACCAACCAACCA  
CCCCGGCCCGGTGCTGTACACCAACCGGCCAGATCATCGCGGACATCCGCAAGGCCCACTGCAACATCTCCAAGAACCAAGTGAACCAACACCC  
TGGGCCAGATCGTGAAGAAGCTGCGCGAGCACTTATGAACAAGACCATCGTGTTCAGCGCTCTCCGGCGGCAACCCCGAGATCGTGATG  
CACTCCTTCAACTGCGCGCGGCGAGTCTTCTACTGCAACCAACCCAGCTGTTCAACTCCACCTGGCGTCCAACTCCACCTGGAACGACAC  
CACCGAGACCAACAACACCGACCTGATCACCTGCCCTGCCGATCAAGAGATCGTGAAACATGTGGCAGAGGTGGGCAAGGCCATGTACG  
CCCCCCCCTTCGGCCAGATCCGCTGCTCCTCCAACATCACCGGCTGCTGCTGATCCGCGACGGCGCTCCAACAACACCGAGACCTTC  
CGCCCCGGCGGCAACATGAAGGACAACTGGCGCTCCGAGCTGTACAAGTACAAGTGTGAAGATCGAGCCCTTGGCGGTGGCCCCC  
CGCGCCCAAGCGCGGTGGTGCAGCGCGAGAGCGCGCTGGGCATCGCGCCCTGCTGTTCCGCTTCTGGCGCGCGCGGCTCCACCA  
TGGCGCGCGCTCCATGACCTGACCGTGCAGGCGCGCGAGCTGCTGTCCGGCATCGTGCAGCAGCAACAACCTGCTGCGCGCCATCGAG  
GCCAGCAGCACATGCTGCAGTGAACCGTGTGGGCGCATCAAGCAGCTGCAGGCCCGCGTGTGGCGCTGAGCGCTACCTGAAGGACCAAGCA  
GCTGTGGGCATCTGGGGCTGCTCCGGCAAGTGTGTCACCAACCGTGGCCCTGGAACGCCCTCCTGGTCCAACAAGTCCCTGGACGACA  
TCTGGAACAACATGACCTGGATGGAGTGGAGCGCGAGATCGACAACATACACGGCTGATCTACACCTGATCGAGCAGTCCCAAGAACCAAG  
CAGGAGCGCAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGCCCTCCCTGTGGAACCTGGTTCAACATCACCAACTGGCTGTGGTACATCAA  
GATCTTCATCATGATCATCGCGGCTGATCGGCTGCGCATCGTGTCCCGTGTCTCATCATCAACCGCGTGCAGAGGGCTACTCCC  
CCCTGTCTCCAGACCTGACCCACCAACGCGAGCCCGGACCGCCCGCGCATCGAGGAGGAGGCGCGGAGCAGGACAGGACCGC  
TCCATCCGCCCTGGTGTCCGGCTTCCCTGGCCCTGGGACGACCTGGCTCCCTGTGCTTCTCTACACCGCTGCGGACTTCAT  
CCTGATCGCGCCCGACCGTGGAGCTGCTGGCCCGCTCCTCCCTGAAGGCCCTGGCCCTGGGCTGGGAGGCCCTGAAGTACCTGTGGAACC  
TGCTGTGTACTGGGGCCCGAGCTGAAGAACTCCGCCATCAACCTGCTGGACACCGTGGCCATCGCCGTGGCCAACTGGACCGGACCGCGCC  
ATCGAGGTGGTGCAGCGCGTGGCGCGCGCTGCTGAACATCCCCGTGCGCATCCGCCAGGGCTGGAGCGCGCCCTGCTGTAA



## Centralized HIV-1 gag/nef/pol Protein and the Codon-optimized Gene Sequences

Fig. 63A

## 1. 2003\_CON\_S\_gag.PEP

MGARASVLGGKLDLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEGCOQIIEQLQPALQTGSEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEEEQNKSKQKTQAAADTGNSSKVSQNYPIVQNLOGMVHQAI SPRTLNAWVKVVEEKAFSPEVIFMFSALSEGATPQDL  
 NTMLNTVGGHQAAQMQLKDTINEEAAEWDRLLHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSILDIRQGPKEPRFDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANPDCKTILKALPGGATLEEMMTACQGVGGPSPSHKARVLAEAMS  
 QVTNTTIMQRGNFKGQKRIIKCFNCGKEGHIARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWPSNKGPRPGNFIQSRPEPTAPPAAE  
 SFGFGEETTPSPKQEPKDKELYPLASLSLFGNDPLSQ\$

Fig. 63B

## 2003\_CON\_S\_gag.OPT

ATGGGCGCCGCGCTCCGTGTCTCGGCGGCAAGCTGGACGCTGGAGAAAGATCCGCTTGGCCCCCGGCGGCAAGAAGTACCGCCT  
 GAAGCACTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGGCTGCCAGCATCATCG  
 AGAGCTGACGCCCGCTGCAGACCGGCTCCGAGGAGTGGCTCCCTGTACAAACCGTGGCCACCTGTACTGGGTGCACCGCGCATC  
 GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAAAGCAAGAACCCAGAGCGCGCGCCGACACCGG  
 CAACTCCTCCAAAGGTGTCCAGAAC TACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGACCAAGGCTATCTCCCGCCACCTCCAGGGCGCCACCCCGCAGGACCTG  
 CCTGGGTGAAGGTGGTGGGCGGCCACCAAGCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGCGCGGAGTGGGACCGCCT  
 GCACCCGTGCACGCGGCCCATCCCCCGGCGAGATGCGGAGCCCGGCTCCGACATCGCCGCAACCTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCTCCAAACCCCGCATCCCCGTGGCGAGATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGCCCCCAAGGAGCCCTTCGCCGACTACGTGGAACCGCTTCTTCAAGACCTGCGGCGCGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCTGTGTGCAGAACGCCAACCCGACTGCAAGACCATCTCTGAAGGCCCTGG  
 GCCCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCCAAGGCCCGGTGTGGCGAGGCCATGTCC  
 CAGGTGACCAACACCAACCATCATGATGCAGCGCGGCAACTTCAAGGCCAGAAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCCA  
 CATCGCCCGCAACTGCCGCGCCCCCGCAAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCCACCAAGATGAAGGACTGCACCGAGGCCAGG  
 CCAACTTCTGGGCAAGATCTGGCCCTCCAAAGGGCGGCCCGGCCAACTTCTGCACTCCGCCCCGAGCCACCGCCCCCGCCGAG  
 TCCTTCGGCTTCGGCGAGGAGATCA CCCCCCTCCCAAGCAGGAGGCCCAAGGAGCTGTACCCCCCTGGCCTCCCTGAAGTCCCTGTT  
 CGGCAACGACCCCCCTGTCCAGTAA

103/178



Fig. 64A

## 2. 2003 M.GROUP.anc gag.PEP

MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLEAEGCQIMQLPALQTGTEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEEQNKSQKTQAAADKGDSSQVSQNYPIVQNQQMVHQAISPRTLNAAVVKVVEEKAFSPVPIPMFSALSEGATPQDL  
 NTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSIIDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANPDKTILKALGPATLEEMMTACQGVGGPGHKARVLAEMS  
 QVTNANIMMQRGNFKPPRRIVKCFNCGKEGHIARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWPSNKGPRGNFLQSRPEPTAPPAE  
 SFGFGEIITPSPKQEPKDKELYPLASLKSIFGSDPLSQ\$

Fig. 64B

## 2003 M.GROUP.anc gag.OPT

104/178

ATGGGCGCCCGCGCTCCGTGTCTCGGGGGCAAGCTGGAGCGCTGGGAGAAGATCCGGCTGCGCCCGCGCGGAAGAAGTACCGCCT  
 GAAGCACCTGCTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCGCCGAGGGTGCACGAGATCATGG  
 GCCAGCTGCAGCCCGCCTGCAGACCGGCACCGAGGAGCTGCGTCCCTGTACAACACCGTGGCCACCCCTGTAAGTGTGACAGCGCATC  
 GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAGTCCAGCAGAAGACCCAGCAGCGCCGCGGACAAAGG  
 CGACTCCTCCAGGTGTCCAGAACTACCCCATCGTGAGAACCTGCAGGCCAGATGGTGACACAGGCCATCTCCCCCGCACCCCTGAACG  
 CCTGGGTGAAGGTGGTGGAGGAGAGGCCTTCTCCCCGAGGTGATCCCCATGTTCTCGCCCTGTCCGAGGGCGCCACCCCGAGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACCGCCCATGCAGATGCTGAAGGACACCATCAACGAGGAGCGCCGAGTGGGACCGCCT  
 GCACCCCGTGCACCGCGCCCATCCCCCGCCAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACCCCTCCACCCCTGCAGGAGC  
 AGATCGGCTGGATGACCTCCAACCCCCCATCCCCGTGGCGAGATCTACAAGCGCTGGATCATCTTGGGCTGAACAAGATCGTGGCATG  
 TACTCCCCCGTGTCCATCTTGACATCCGCCAGGGCCCCAAGAGCCCTTCGCGGACTACGTGGACCGCTTCTCAAGACCCCTGCGCGCCGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGAAGGCCCTGG  
 GCCCGGCGCCACCCCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGGCCACAAGGCCCGCTGTGGCCGAGGCCATGTCC  
 CAGGTGACCAACGCCAACATCATGATGACGCGCGCAACTCAAGGGCCCCCGCGCATCTGTGAAGTCTTCAACTGCGGCAAGGAGGCCA  
 CATCGCCGCAACTGCGCGGCCCGCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCATGAAGGACTGCACCGAGCGCCAGG  
 CCAACTTCTGGGCAAGATCTGGCCCTTCCAACAAGGGCGCCCGGCAACTTCTGAGTCCCGCCCGAGCCCAAGCCCCCGCCGAG  
 TCCTTCGGCTTCGGCGAGGAGATCACCCCTCCCCCAAGCAGGAGGCCCAAGGAGCTGTACCCCTGGCCCTCCCTGAAGTCCCTGTT  
 CGGCTCCGACCCCTGTCCCCAGTAA

Fig. 65A

## 3. 2003 CON A1 gag.PEP

MGARASVLGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPSSLLETTGCGQQIMEQLPALKTGTEELRSLYNTVATLYCVHQRI  
 DVKDTKEALDKIEIQNKSQKTQQAADTGNSSKVSQNYPIVQNAQGMVHQSLSPRTINAWVKVIEEKAFSPEVIPMFSALSEGATPQDL  
 NMMLNIVGGHQAAQMLKDTINEEAAEWDRLHPVHAGPIPPGQMRPRGSDIAGTTSTPQEQIGWMTGNPPIPVGDIYKRWIILGLNKIVRM  
 YSPVSILDIKQGPKEPRFDYVDRFFKTLRAEQATQEVKNWMTETLLVQANPDCKSILRALPGATLEEMMTACQGVGGPGHKARVLAEAMS  
 QVQHTNIMMQRGNFRGQRIKCFNCGKEGHLARNCRAPRKKGWKCGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFPQSRPEPTAPPAEI  
 FGMGEIITSPPKQEQKDREQDPPPLVLSLFGNDPLSQ\$

Fig. 65B

## 3. 2003 CON A1 gag.OPT

ATGGGCGCCCGCCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCCTGCGCCCCGGCGGCAAGAAGATACCGCCT  
 GAAGCACCTGGTGTGGCCCTCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCAACCGAGGGTGCCAGAGATCATGG  
 AGCAGCTGCAGCCCGCCCTGAAGACCGGCAACCGAGAGCTGGCTCCCTGTACAAACACCGTGGCCACCTGTACTGCTGCACCAAGCGCATC  
 GACGTGAAGGACACCAAGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAAGCAGAAGACCAAGCGCCGCCGACACCGG  
 CAACTCCTCCAAGGTGTCCAGAACTACCCCATCGTGCAAGAACGCCAGGCCAGATGGTGCAACAGTCCCTGT'CCCCCGCACCCCTGAACG  
 CCTGGGTGAAGGTGATCGAGGAGAAGGCCTT'CCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGGCGGCCACAGGCCGCGCATGCAGATGCTGAAGGACACCAATCAAGAGAGGCCCGCGAGTGGACCCCT  
 GCACCCCGTGCAAGCGGCCCTCCATCCCGCGCGAGATGCGGAGCCCGCGGCTCCGACATCGCCGCAACCACTCCACCCCGCAGGAGC  
 AGATCGGCTGGATGACCGGCAACCCCGCATCCCGGTGGCGACATCTACAAGCGTGGATCATCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCCGTGTCATCCTGGACATCAAGCAGGCCCCCAAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCTGCGCGCCGA  
 GCAGGCCACCCAGGAGTGAAGAACTGGATGACCGGACCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGCGCCCTGG  
 GCCCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGCGTGGCGGCCCGCCAGCCCGCGTGTGCTGGCCGAGGCCATGTCC  
 CAGGTGCAGCACACCAACATCATGATGCAGCGCGCAACTTCCGCGGCCAAGCGCATCAAGTGTCTCAACTGGGCAAGGAGGCCACCT  
 GGCCCGCAACTGCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGGACTGCACCGAGCGCCAGGCCA  
 ACTTCTGGGCAAGATCTGCCCTCCTCCAAGGGCGGCCCGGCAACTTCCCCAGTCCCGCCCCGAGCCACCGCCCCCGCGGAGATC  
 TTCGGCATGGCGGAGGAGATCACCTCCCCCCCCAAGCAGGAGCAGAACCGGAGACCCCCCCCCCTGGTGTCCCTGAAGTCCCTTGTT  
 CGGCAACGACCCCTGTCCCCAGTAA

105/178

Fig. 65C

## 4. 2003 A1.anc gag.PEP

MGARASVLSGGKLD AWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETAEGCQIQMQLPALKTGTEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEIIONKSKQKTQAAADTGNSSKVSQNYPIVQNAQGMVHQSLSPRTLNAWVKVIEEKAFSPEVIMFSAISEGATPQDL  
 NMMLNIVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTGNPPPIPVGDIYKRWIILGLNKIVRM  
 YSPVSLDIRQGPKEFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQNNANPDCKSILRALGPGATLEEMMTACQGVGGPGHKARVLAEAMS  
 QVQNTDIMMQRGNFRGPKRIKCFNCGKEGHLARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFPQSRPEPTAPPAEN  
 FGMGEEMISSPKQEQKDREQYPPLVSLKSLFGNDPLSQ\$

Fig. 65D

## 2003 A1.anc gag.OPT

106/178

ATGGGCGCCCGCGCTCCGTGCTGTCCGGCGGCAAGCTGGACGCGCTGGGAGAAGATCCGCGCTGCGCCCCGGCGGCAAGAAAGTACCGCCT  
 GAAGCACCTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCGCTGAACCCCGGCTGTCTGGAGACCGCGGAGGCTGCCAGCATCATGG  
 GCCAGCTGACGCGCCCTGAAGACCGGCAACCGAGGAGCTCGGCTCCCTGTACAACACCGTGGCCACCTGTACTGCGTGCCAGCGCATC  
 GAGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAAGTCCAGCAGAAAGCCAGAGGCCCGCGCGACACCGG  
 CACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACGCCAGGCCAGATGTTGCCACCAAGTCCCTGTCCCCCGCACCTGAACG  
 CCTGGGTGAAGGTGATCGAGGAGAGGCTTCTCCCCGAGGTGATCCCCATGTTCTCGCGCTGTCCGAGGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACCAAGCCGCGCATGACAGTGTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGGACCGCCT  
 GCACCCCGTGACGCGGCCCATCCCCCGGCGAGATGCGCGAGCCCCGCGGTCCGACATCGCCGGCACCCACCTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCGGCAACCCCGCATCCCCGTGGCGACATCTACAAGCGCTGGATCATCTCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCCGTGTCCATCTGGACATCCGCGAGGCGCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTCAAGACCTGCGCGCCGA  
 GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCTGTGTGTGAGAACGCCAACCCCGACTGCAAGTCCATCTGCGCGCCCTGG  
 GCGCGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGCGTGGCGGCGCCCGGCCACAAAGGCCCGGTGCTGGCCGAGGCCATGTCC  
 CAGGTGCAGAACACCGACATCATGATGCAGCGGCAACTTCGCGGCCCCAAGCGCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACCT  
 GCGCGCAACTGCGCGCCCCCGCAAGAGGCTGTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGACTGCACCGAGCGCCAGGCCA  
 ACTTCTGGGCAAGATCTGGCCCTCTTCAAGGCGCGCCCGCAACTTCCCCAGTCCCCCGCCGAGCCCCACCGCCCCCGCGGAGAAC  
 TTCGGCATGGCGGAGGAGATGATCTCTTCCCCCAAGCAGGAGCAGGACCGCGGAGCAGTACCCCCCTGGTGTCCCTGAAGTCCCTGTT  
 CGGCAACGACCCCTGTCCCAGTAA

Fig. 66A

## 5. 2003\_CON A2 gag.PEP

MGARASILSGGKLDWEKIRLRPGGKKYRLKHLVWASRELEKFSINPSLLETSEGRQIIROLQALQGTTELKSLYNTVAVLVCVHQRI  
 DVKDTKEALDKIEEQNKCKQKTQHAAADTGNSSSSQNYPIVQNAQGMVHQAI SPRTLNAWVKVVEEKA FSPEVIPMFTALSEGATPQDL  
 NTMLNTVGGHQAAMQLKDTINEEAAEWDLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQANPDCKSILRALPGATLEEMMTACQGVGGPSHKARVLAEAMS  
 QVQNTNTNIMMQRGNFRGQKRIKFCNCGKEGHLARNCRAPRKKGCKGEGHMKDCTERQANFLGKIWPSNKGPRGNFPQSRTEPTAPPA  
 ENLRMGEEITSSLKQELKTREYPNPAISLKLFGNDPLSQ\$

Fig. 66B

## 2003\_CON A2 gag.OPT

ATGGGCGCCGCGCCTCCATCCTGTCCGGCGGCAAGCTGGACGCCCTGGGAGAAAGATCCGCCCTGGCCCCGGGCAAGAAGTACCGCCT  
 GAAGCACCTGGTGTGGCCTCCCGGAGCTGGAGAAAGTTCTCATCAACCCCTCCCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCATCC  
 GCCAGCTGCAGCCCGCCTGCAGACCGGCACCGAGGAGCTGAAGTCCCTGTACAAACCGTGCCCGTGTGTACTGCGTGACACCGCATC  
 GACGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAAGTGCAAGCAGAACAGACCCAGCACGCCGCCGCCGACACCGG  
 CAATCTCTCTCTCTCCAGAACTACCCCATCGTGACAGAACGCCAGGCCAGATGGTGACACCGAGCCATCTCCCCCGCACCCCTGAACG  
 CCTGGGTGAAGTGTGGAGGAGAGCCCTTCTCCCCGAGGTGATCCCATGTTCACCGCCCTGTCCGAGGGCGCCACCCCCAGGACCTG  
 AACACCATGCTGAACACCGTGGGCGGCCACAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGGACCCGCT  
 GCACCCCGTGCACGCCGGCCCCATCCCCCGGCCAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCTCCAAACCCCCCATCCCCGTGGCGAGATCTACAAGCGCTGGATCATCCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCCGTGTCCATCCTGGACATCCGCCAGGCCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACCTGCCGCCCGA  
 GCAGGCCACCCAGGAGTGAAGAACTGGATGACCGACACCCCTGCTGTGCAGAACGCCAACCCGACTGCAAGTCCATCCTGGCGGCCCTGG  
 GCCCCGGCGCCACCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCTCCCAAGGCCCGCGTGTGCTGGCCGAGGCCATGTCC  
 CAGGTGCAGAACACCAACACATCATGATGCAGCGCGGCAACTTCCGCGGCCAGAACGCGCATCAAGTGTTCAACTGGGGCAAGGAGGG  
 CCACCTGGCCCGCAACTGCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGCGCAAGGAGGCCACCATCAAGTGAAGGACTGCACCGAGCGCC  
 AGGCCAACTTCTTGGGCAAGATCTGGCCCTCCAAACAAGGGCCCCCGGCAACTTCCCCAGTCCCGCACCGAGCCACCGCCCCCGGCC  
 GAGAACTGCGCATGGCGGAGGAGATCACCTCCTCCCTGAAGCAGGAGCTGAAGACCCCGGAGCCCTACAACCCCGCCATCTCCCTGAAGTC  
 CCTGTTCGGCAACGACCCCTGTCCCCAGTAA

Fig. 67A

## 6. 2003 CON B gag. PEP

MGARASVLSGGELDRWEKIRLRPGGKKKYLKHIVWASRELERFAVNPGLLETSEGRQILQLPSLQTSSEELRSLYNTVATLYCVHQRI  
 EVKDTKEALEKIEEEQNKSKKAQAAADTGNSSQVSONYPIVQNLQGMVHQAI SPRTLNAWVKVVEEKAFSPVIMFSA LSEGATPQDL  
 NTMLNTVGGHQAAQMMLKETINEEAAEWDRLHPVHAGPIAPGQMRPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIIILGLNKIVRM  
 YSPTSILDIRQPEKPEFRDYVDREYKTLRAEQASQEVKNWMTETLLVQANANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMS  
 QVNTSATIMMQRGNFRNQRTVKCFNCGKEGHIAKNCRAPRKKGCWKCKGHEGHQMKDCTERQANFLGKIWP SHKGRPGNFLOSRPEPTAPPE  
 ESFRFGEETTPSQKQEPIDKELYPLAS\$

Fig. 67B

## 2003 CON B gag. OPT

ATGGGCGCGCGCCTCCGTGTCGCGGGCGGAGCTGGACCGCTGGGAGAAGATCCGCCCTGCGCCCCGCGGCAAGAAAGTACAAAGCT  
 GAAGCACATCGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCGTGAACCCCGGCCCTGCTGGAGACCTCCGAGGGCTGCCGCGAGATCCTTGG  
 GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTCGCTCCCTGTACAACACCGTGCCACCCCTGTACTGCGTGCACCGCGCATC  
 GAGGTGAAGGACACCAAGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAGTCCAGAAGAAGGCCCGCAGAGGCCCGCGCACACCGG  
 CAACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGACACGAGCCATCTCCCCCGCACCTGAACG  
 CCTGGTGAAGGTGGTGGAGAGGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCCACCCCGCAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACCAAGCCGCCATGCAGATGCTGAGGAGACCATCAACGAGGAGGCCCGCGAGTGGGACCCCT  
 GCACCCGTGCACGCCGCCCATCGCCCCCGGCCAGATCGCGAGCCCCCGGCTCCGACATCGCCGGCACCATCTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCAACACCCCCCATCCCCGTGGCGGAGATCTACAAGCGCTGGATCATCTTGGGCCCTGAACAAGATCGTCCGCATG  
 TACTCCCCACCTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCCGGACTACGTGGACCCGCTTCTACAAGACCTTGCAGGCCG  
 GCAGGCCCTCCAGAGGTGAAGAACTGGATGACCGAGACCTGTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGAAGGCCCTGG  
 GCCCGCGGCCACCTTGAGGAGATGATGACCGCTGCCAGGGGTGGCGGCGGCCCGCCACAAGGCCCGCGTGTGGCCGAGGCCATGTCC  
 CAGGTGACCAACTCCGCCACCATCATGATGACGCGGGCAACTTCCGCAACCAAGCAAGACCGTGAAGTCTCAACTGCGGCAAGGAGG  
 CCACATCGCCAAGAACTCGCGGCCCGCCGCAAGAAAGGCTGCTGGAAGTGGGCAAGGAGGCCACCATAGATGAAGGACTGCACCGAGCGCC  
 AGGCCAACTTCTGGGCAAGATCTGGCCCTCCCAACAAGGGCGGCCCGGCAACTTCTGTGAGTCCCGCCCCGAGCCACCGCCCCCGGAG  
 GAGTCCTTCCGCTTCGGCGAGGAGACCAACCCCTCCCAAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCTTGGCCTTCCTAA

108/178

## Fig. 67C

7. 2003 B. *anc* gag. PEP

MGARASVLSGGKLDKWEKIRLRPGGKKYKLLKHIVWASRELERFAVNPGLLETSEGRQILGQLPALQGTSEELRSLYNTVATLYCVHQRI  
 EVKDTKEALDKIEEQNKSKKAQQAADTGNSSQVSQNYPIVQNLQGMVHQAI SPRTLNAWKVVEEKAFSPEVIMFSAISEGATPQDL  
 NTMLNTVGGHQAAQMMLKETINEEAAEWDRLHPVHAGPIAPQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM  
 YSPISILDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNPDPCKTILKALGPAATLEEMMTACQVGGPGHKARVLAEAMS  
 QVTNSTTIMMQRGNFRDQKIVKCFNCGKEGHIARNCRAPRKKGCKGEGHQMCKDCTERQANFLKGIWPSHKGRPGNLFQSRPEPTAPPE  
 ESFRFGEETTPSQKEPIDKELYPLASLKSIFGNDPSSQ\$

## Fig. 67D

2003 B. *anc* gag. OPT

ATGGGCGCCCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGCCCTGCCGCCCGGGCAAGAAGTACAAGCT  
 GAAGCACATCGTGTGGGCTCCCGCGAGCTGGAGCGCTTCGCCGTGAACCCCGGCTGCTGGAGACCTCCGAGGGTGCAGGATCCTGG  
 GCCAGCTGCAGCCCGCCTGCAGACCGGCTCCGAGGAGCTGCCCTCCCTGTACAAACCGTGGCCACCTGTACTGCTGCACACGCGCATC  
 GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAAGTCCAAGAAGAGGCCAGCGGCCCGCGGACACCGG  
 CAACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAAGAACCTGCAGGCCAGATGGTGCACCAAGGCCATCTCCCCCGCACCTGAACG  
 CCTGGGTGAAGTGTGGAGGAGAAGCCCTTCTCCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCAGGACCTG  
 AACACCATGCTGAACACCGTGGGCGGCCACCAAGCCGCCATGCAGATGCTGAAGGAGAGACCATCAACGAGGAGGCCCGCCGAGTGGACCTG  
 GCACCCCGTGCAACGCGGCCCTCCATCGCCCGCAGATCGCGAGCCCGCGGCTCCGACATCGCCGACACCACTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCAACAACCCCCCATCCCGCTGGCGGAGATCTACAAGCGCTGGATCATCTTGGGCTGAACAAGATCGTGGCATG  
 TACTCCCCATCTCCATCTTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTACAAGACCTTGGCGGCCGA  
 GCAGGCTTCCAGGACGTGAAGACTGGATGACCGGACCTGTGGTGCAGAACGCCAACCCGACTGCAAGACCATCTGAAGGCCCTGG  
 GCCCCGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGCCACAAGGCCCGCTGCTGGCCGAGGCCATGTCC  
 CAGGTGACCAACTCCACCATCATGATGCAAGCGGCAACTTCCGCGACCAAGGCAAGATCGTGAAGTCTCAACTGCGGCAAGGAGG  
 CCACATCGCCCGCAACTGCCCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGGACTGCAACGAGCGCC  
 AGCCCACTTCTGGGCAAGATCTGGCCCTCCCAAGGGCGGCCCGGCAACTTCTGCAGTCCCGCCCGAGCCCAACGCCCCCGGAG  
 GAGTCTTCCGCTTCGGCGAGGAGACCAACCCCTCCAGAGCAGGAGCCCATCGACAAGGAGCTGTACCCCTGGCCTCCCTGAAGTC  
 CCTGTTCGGCAACGACCCCTCCTCCAGTAA

109/178

Fig. 68A

8. 2003 CON C gag. PEP  
 MGRASILRGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIIKQLPALQGTTEELRSLYNTVATLYCVHEKI  
 EVRDTKEALDKIEEENKSOQKTQAKAADGKVSQNYPIVQNLQOMVHOAISPTLNAWKVIEEKAFSPVIMFTALSEGATPDQDLNTM  
 INTVGGHQAAAMQMLKDTINEEAAEWDRLHPVHAGPIAPGQMRPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIIILGNKIVRMYS  
 VSILDIKQGPKEFRDYVDRFFKTLRAEQATQDVKNWMTDILLVQNPANPDCKTILRALPGATLEEMMTACQGVGSPSHKARVLAEAMSQAN  
 NTNIMQRSNFKGPKRIVKFCNCGKEGHIARNRAPRKKGCKGEGHQMCKDCTERQANFLGKIWP SHKGRPGNLFQNRPEPTAPAESFR  
 FEETTPAPKQEPKDRPLETSLKSLFGSDPLSQ\$

Fig. 68B

2003 CON C gag. OPT  
 ATGGGCGCCCGCGCCTCCATCCTGCGGGGGCAAGCTGGACAAGTGGGAGAAGATCCGCCCTGGGCCCCGGGGCAAGCACTACATGCT  
 GAAGCACCTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA  
 AGCAGCTGCAGCCCGCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCCGTGGCCACCCCTGTACTGCGTGACGAGAAGATC  
 GAGGTGCGGACACCAAGGAGCCCTGGACAAAGATCGAGGAGGAGCAACAAGTCCAGCAGAAGACCCAGAGGCCAAGGCCCGCCGACGG  
 CAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCAACAGGCCATCTCCCCCGCACCCCTGAACGCCCTGGGTGA  
 AGGTGATCGAGGAGAAAGGCTTCTCCCCGAGGTGATCCCCATGTTACCCGCTGTCCGAGGGCGCCACCCCGAGGACCTGAACACCATG  
 CTGAACACCGTGGGCGGCCACAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGACCGCTGCACCCCGT  
 GCACGCCGCCCCCATCGCCCCCGCCAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACCACTCCACCTCCACCTGCAGGAGCAGATCGCCT  
 GGATGACCTCCAACCCCCCATCCCGTGGCGGACATCAACAGCGCTGGATCATCCTGGGCTGAACAAGATCGTGGCATGTACTCCCC  
 GTGTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCGCGACTACGTGGACCGCTTCTTCAAGACCCCTGCGCGCCCTGGGCCCCGGCG  
 CCAGGACGTGAAGAACTGGATGACCGACACCTGCTGGTGAGAACGCCAACCCGACTGCAAGACCATCTGCGCGCCCTGCGGCCCCGCGG  
 CCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGGCGGCCCTCCACAAAGGCCCGGTGCTGGCCGAGGCCATGTCCAGGCCAAC  
 AACACCAACATCATGATGACGCGCTCCAACCTCAAGGGCCCAAGCGCATCGTGAAGTCTCAACTGCGGCAAGGAGGCCACATCGCCCG  
 CAACTGCCGCGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACACAGATGAAGGACTGCACCGAGGCCAGGCCAACTCC  
 TGGGCAAGATCTGGCCCTCCCAAGGGCGCCCCCGCAACTCCTGCAGAACCGCCCCGAGCCCAAGCCCCCCCCCGAGTCCCTTCGCG  
 TTCGAGGAGACACCCCCGCCCCCAAGCAGGAGGCCCAAGGACCGGAGCCCCCTGACCTCCCTGAAGTCCCTGTTGGCTCCGACCCCCCTGTC  
 CCAAGTAA

110/178

Fig. 68C

## 9. 2003 C.anc.gag.pEP

MGARASILRGKLDITWEKIRLRPGGKKHYMIKHLVWASRELERFALNPGLLETSEGCKQIMKQLPALQQTGEELRSLYNTVATLYCVHERI  
 EVRDTKEALDKITEEEQNKSQKTOQAEEADGNGKVSQNYPIVQNLOGMVHQAISPRTLNAWKVVEEKAFSPVIMFTALSEGATPQDL  
 NTMLNTVGGHQAAQMQLKDTINEEAAEWDRLHPVHAGVPAGQMPREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRM  
 YSPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANPDCKTILRALPGATLEEMMTACQGVGPGHKARVLAEAMS  
 QANNTNIMMQRSNFQPKRIVKCFNCGKEGHIARNCRAPRKKGWCWKGEGHQMKDCTERQANFLGIWPSHKGRPNFLQSRPEPTAPPAE  
 SFRFEETTPAPKQEPKPREPLTSLKSLFGSDPLSQ\$

Fig. 68D

## 2003 C.anc.gag.OPT

ATGGGCGCGCGCCCTCCATCCTGCGCGCGCGGCAAGCTGGACACCTGGAGAAGATCCGCTGCGCCCGCGGCAAGACCACTACATGAT  
 CAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATGA  
 AGCAGCTGCAGCCCGCCCTGCAGACCGGACCGGAGGCTGCGCTCCCTGTACAACACCCGTGGCCACCTGTACTGCTGCACGAGCGCATC  
 GAGGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGAGAACAAAGTCCAGCAGAAGACCCAGCAGGCCGAGGCCGCGACGG  
 CGACAACGGCAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGCTGCACCCAGGCCATCTCCCCCGCACCTGAACG  
 CCTGGGTGAAGGTGGTGGAGGAGAAAGGCTTCTCCCCGAGGTGATCCCCATGTTCAACCGCCCTGTCCGAGGGCGCACCCCGAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACAGCGGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCAGTGGGACCGCCT  
 GCACCCCGTGCAAGCGCGCCCGTGGCCCCCGGCGAGTGGCGAGCCCCCGGCTCCGACATCGCCGCGCACCATCTCAACCTGCAGGAGC  
 AGATCGCCTGGATGACCTCCAACCCCATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGGACCGCTTCTCAAGACCTGCGCGCCGA  
 TACTCCCCGTGTCCATCCTGGACATCAAGCAGGCGCCCAAGAGCCCTTCGCGACTACGTGGACCGCTTCTCAAGACCTGCGCGCCGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCTGCGCGCCCTGG  
 GCGCGCGCCCAACCTGGAGGAGATGATGACCGCCTGCCAGGCGCTGGCGGCGCCCGGCCACAAAGCCCGCTGCTGGCCGAGGCCATGTCC  
 CAGGCCAACAAACCAACATCATGATGACGCGCTCCAACCTCAAGGGCCCCAAGCGCATCGTGAAGTGTCAACTGCGGCAAGGAGGGCCA  
 CATCGCCCGCAACTGCCGCGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGACTGCACCGAGCGCCAGG  
 CCAACTTCCCTGGGCAAGATCTGGCCCTCCCAACAAGGGCGCCCCCGGCAACTTCTGCAAGTCCCGCCCGAGCCCAACCGCCCCCGCGAG  
 TCCTTCCGCTTCGAGGAGACCAACCCCGCCCCCAAGCAGGAGCCCAAGGACCGGAGCCCTGACCTCCCTGAAGTCCCTGTTCCGCTCCGA  
 CCCCCGTGCCAGTAA

111/178



Fig. 69A

## 10. 2003\_CON D gag .PEP

MGARASVLSGGKLD<sup>1</sup>AW<sup>2</sup>EKIRL<sup>3</sup>RP<sup>4</sup>GGKKYRL<sup>5</sup>KHIV<sup>6</sup>ASRELERFALN<sup>7</sup>PGLLE<sup>8</sup>TSE<sup>9</sup>GCKQI<sup>10</sup>IG<sup>11</sup>QL<sup>12</sup>QPAI<sup>13</sup>QT<sup>14</sup>SEEL<sup>15</sup>RL<sup>16</sup>SYNT<sup>17</sup>VATLY<sup>18</sup>CV<sup>19</sup>HERI<sup>20</sup>  
 EVKDTKEALEKIEE<sup>21</sup>EQNKS<sup>22</sup>KKKAQ<sup>23</sup>AAAD<sup>24</sup>TGN<sup>25</sup>SSQ<sup>26</sup>VSQ<sup>27</sup>NPY<sup>28</sup>IV<sup>29</sup>QNL<sup>30</sup>QMV<sup>31</sup>HQAI<sup>32</sup>SP<sup>33</sup>RTL<sup>34</sup>NAW<sup>35</sup>KV<sup>36</sup>IEE<sup>37</sup>KA<sup>38</sup>FS<sup>39</sup>PE<sup>40</sup>VIP<sup>41</sup>MF<sup>42</sup>SAL<sup>43</sup>SEGAT<sup>44</sup>PQDL<sup>45</sup>  
 NTMLNTVGGHQAAMQ<sup>46</sup>MLKET<sup>47</sup>INEEAAEW<sup>48</sup>DR<sup>49</sup>LHPV<sup>50</sup>HAG<sup>51</sup>VPAG<sup>52</sup>QMR<sup>53</sup>PRGSDI<sup>54</sup>AG<sup>55</sup>TT<sup>56</sup>SL<sup>57</sup>QEQ<sup>58</sup>IG<sup>59</sup>WMT<sup>60</sup>SN<sup>61</sup>PP<sup>62</sup>IP<sup>63</sup>VGEI<sup>64</sup>YK<sup>65</sup>RW<sup>66</sup>IIL<sup>67</sup>GLN<sup>68</sup>KI<sup>69</sup>VRM<sup>70</sup>  
 YSPV<sup>71</sup>SILDIR<sup>72</sup>QGPKE<sup>73</sup>PF<sup>74</sup>RDY<sup>75</sup>VDR<sup>76</sup>FYK<sup>77</sup>TLRAE<sup>78</sup>QASQ<sup>79</sup>DV<sup>80</sup>KNW<sup>81</sup>MTET<sup>82</sup>LL<sup>83</sup>VQ<sup>84</sup>NAN<sup>85</sup>PD<sup>86</sup>CK<sup>87</sup>TI<sup>88</sup>LKAL<sup>89</sup>GP<sup>90</sup>EAT<sup>91</sup>LEEM<sup>92</sup>MTAC<sup>93</sup>QGV<sup>94</sup>GPS<sup>95</sup>HKAR<sup>96</sup>VLA<sup>97</sup>EAMS<sup>98</sup>  
 QATNSAAVM<sup>99</sup>Q<sup>100</sup>RGN<sup>101</sup>FK<sup>102</sup>PR<sup>103</sup>KI<sup>104</sup>IK<sup>105</sup>FN<sup>106</sup>CG<sup>107</sup>KEGH<sup>108</sup>IA<sup>109</sup>KN<sup>110</sup>CR<sup>111</sup>PK<sup>112</sup>KG<sup>113</sup>CK<sup>114</sup>W<sup>115</sup>CG<sup>116</sup>KEGH<sup>117</sup>Q<sup>118</sup>MD<sup>119</sup>CTER<sup>120</sup>QAN<sup>121</sup>FL<sup>122</sup>GKI<sup>123</sup>W<sup>124</sup>PS<sup>125</sup>HK<sup>126</sup>RP<sup>127</sup>GN<sup>128</sup>FL<sup>129</sup>Q<sup>130</sup>SR<sup>131</sup>PE<sup>132</sup>PTAP<sup>133</sup>PA  
 ESFGFGEI<sup>134</sup>TPSQ<sup>135</sup>KQEQ<sup>136</sup>KDKELY<sup>137</sup>PL<sup>138</sup>TS<sup>139</sup>LK<sup>140</sup>SL<sup>141</sup>FG<sup>142</sup>ND<sup>143</sup>PL<sup>144</sup>SSQ<sup>145</sup>

Fig. 69B

## 2003\_CON D gag .OPT

ATGGCGCCCGCGCCT<sup>1</sup>CCGT<sup>2</sup>CGT<sup>3</sup>GTCCGGCGGCAAG<sup>4</sup>CTGGACGCT<sup>5</sup>GGGAGAAAGAT<sup>6</sup>CCGCCCT<sup>7</sup>GGCGCCCCCGGGCGCAAGAAGAT<sup>8</sup>ACCGCCT<sup>9</sup>  
 GAAGACAT<sup>10</sup>CGTGTGGGCT<sup>11</sup>CCCGGAGCTGGAGCGCT<sup>12</sup>TCGCCCT<sup>13</sup>GAAACCCCGGCT<sup>14</sup>GCTGGAGACCT<sup>15</sup>CCGAGGGCTGCAAGCAGAT<sup>16</sup>CATCG<sup>17</sup>  
 GCCAGTGCAGCCCGCAT<sup>18</sup>CCAGACCGGCTCCGAGGAGCT<sup>19</sup>CGCTCCCTGTACAAACACCCGTGGCCACCTGTACTGCGTGCACGAGCGCATC<sup>20</sup>  
 GAGGTGAAGGACACCAAGGAGGCCCTGGAGAAAGAT<sup>21</sup>CGAGGAGGAGCAGAACTCCAAAGAAAGGCCACAGCAGCGCCGCGCCGACACCGG<sup>22</sup>  
 CAACTCCTCCAGGTGCCAGAACTACCCCAT<sup>23</sup>CGTGCAAGAACCTGCAGGGCCAGATGGTGACAGGCCAT<sup>24</sup>TCCCCCGCACCCCTGAAAC<sup>25</sup>  
 CCTGGTGAAGTGATCGAGGAGAAGCCCTTCTCCCCGAGGTGATCCCCAT<sup>26</sup>TTCTCCGCCCTGTCCGAGGGCGCCACCCCAAGGACCTG<sup>27</sup>  
 AACACCATGCTGAACACCCGTGGCGGCCACACAGGCCGCCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT<sup>28</sup>  
 GCACCCCGTGACGCCGCCCGTGGCCCCCGGCCAGATGCGCGAGGCCCGCGGCTCCGACATCGCCGACACCATCCACCTCCACCTGCAGGAGC<sup>29</sup>  
 AGATCGGCTGGATGACCTCCAACCCCCCATCCCCCGTGGCGAGATCTACAAGCGCTGGATCATCCTGGGCTGAACAAGATCGTGCGCATG<sup>30</sup>  
 TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACCCCTGCGCGCCGA<sup>31</sup>  
 GCAGGCCCTCCAGGACGTGAAGAACTGGATGACCGAGACCTGTGTTGTCAGAACGCCAACCCGACTGCAAGACCATCCTGAAGGCCCTGG<sup>32</sup>  
 GCCCCGAGGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCACAAAGGCCCGCGTGTGGCCGAGGCCATGTCC<sup>33</sup>  
 CAGGCCACCAACTCCGCCCGCGTGTGATGATGACGCGCGCAACTCAAGGGCCCCCGCAAGATCATCAAGTGTCTCAACTGCGGCAAGGAGG<sup>34</sup>  
 CCACATCGCCAAGAACTCCCGCGCCCCCGCAAGAGGGCTGTGGAAGTGCAGGAGGGCCACCAAGATGAAGGACTGCACCGAGCGCC<sup>35</sup>  
 AGGCCAACTTCTTGGGCAAGATCTGGCCCTCCCAACAAGGGCCCCCGGCAACTTCTCTGAGTCCCCGCCCGAGCCACCGCCCCCCCCCGCC<sup>36</sup>  
 GAGTCTTCCGCTTCCGGCGAGGAGATCACCCCCCTCCCAGAAGCAGGAGCAGAGGAGCTGTACCCCCCTGACCTCCCTGAAGTCCCT<sup>37</sup>  
 GTTCGGCAACGACCCCCCTGTCCCAGTAA<sup>38</sup>

Fig. 70A

## 11. 2003 CON F gag. PEP

MGARASVL<sup>~</sup>SGGKLD<sup>~</sup>DAWEKIRLRPGGKKYRMKHLVWASRELERFALDPGLLETSEGCQKIIGQLQPSLQTSSEELRSLYNTVAVL<sup>~</sup>CVHQKV  
 EVKDTKEALEKLEEEQNK<sup>~</sup>SQKTQAAADKGV<sup>~</sup>SNYP<sup>~</sup>IVQNLQGMVHQAISPRTLN<sup>~</sup>AVKVIEEKAFSP<sup>~</sup>VI<sup>~</sup>PMFSALSEGATPQDLN<sup>~</sup>TML  
 NTVGGHQAA<sup>~</sup>MQMLKDTINEEAAEWDR<sup>~</sup>LHPVHAGPIPPGQMREPRGSDIAGTSTLQEQI<sup>~</sup>QWMTSNPPVPVGD<sup>~</sup>IYKRWII<sup>~</sup>LGINKIVRMYS<sup>~</sup>SPV  
 SILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQANPDC<sup>~</sup>TKILKALPGGATLEEMMTACQGVGGPGHKARVLAEAMSQATN  
 TAIMMQKS<sup>~</sup>NFKQRRIVKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSNKG<sup>~</sup>RPGN<sup>~</sup>FLQSRPEPTAPPAESFGF  
 REEITPSPKQEQKDEGLYPPLASLSLFGNDP\$

Fig. 70B

## 2003 CON F gag. OPT

ATGGGCGCCGCGCCTCCGTGTCCGGGGGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGCGCCCCGGCGGCAAGAAGATACCGCAT  
 GAAGCACCTGGTGTGGCCTCCCGAGCTGGAGCGCTTCGCCCTGGACCCCGGCTGCTGGAGACCTCCGAGGGCTGCCAGAAGATCATCG  
 GCCAGTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCGTGCTGTACTGCGTGCA<sup>~</sup>CCAGAAAGTG  
 GAGGTGAAGGACACCAAGGAGGCCCTGGAGAAGCTGGAGGAGGAGCAGAACAAGTCCACAGAGAAGACCCAGAGGCCGCCGCCGACAAAGG  
 CGTGTCCAGAACTACCCCATCGTGCAGAACTGCAGGGCCAGATGGTGACCAAGGCCATCTCCCCCGCACCCCTGAACGCTGGGTGAAGG  
 TGATCGAGGAGAAGGCTTCTCCCCGAGGTGATCCCCATGTTCTCGCCCTGTCCGAGGGCGCACCCCGAGGACCTGAACACCATGCTG  
 AACACCGTGGCGGCCACAGGCCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGGAGTGGGACCGCTGCACCCCGTGCA  
 CGCCGCCCATCCCCCGGCCAGATGCGGAGCCCCGGGCTCCGACATCGCCGGCACCCACCTCCACCTGCAGGAGCAGATCCAGTGGA  
 TGACCTCCAACCCCCCGTGCCTGGCGGACATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGCGCATGTACTCCCCCGTG  
 TCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCTTGGCGCCGAGAGGCCACCCA  
 GGAGGTGAAGGGCTGGATGACCGACACCCCTGCTGGTGCAAGCCCAACAGCCCGCGTGTGGCCGAGGCCATGTCCCAGGCCACCAAC  
 CCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGGCCACAAGCCCGCGTGTGGCCGAGGCCATGTCCCAGGCCACCAAC  
 ACCGCCATCATGATGCAGAA<sup>~</sup>GTCCAACTTCAAGGGCCAGCGCCGCTCGTGAAGTGTCAACTGCGGCAAGGAGGCCACATCGCCCAAGAA  
 CTGCCGCGCCCCCGCAAGAAGGGCTGCTGAAGTGGCGCGGAGGCCACCAAGTGAAGGACTGCACGAGCGCCAGGCCAACTTCCCTGG  
 GCAAGATCTGGCCCTCCCAACAAGGGCGCCCCGGCAACTTCTGTGAGTCCCGCCCCGAGCCACCGCCCCCGCGAGTCTTCGGCTTC  
 CGCGAGGAGATCACCCCTCCCCCAAGCAGGAGCAGAAGGACGAGGGCTGTACCCCCCTGGCCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCTAA

113/178

114/178

Fig. 71A

**12. 2003 CON G gag. PEP**

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

**Fig. 71B**

2003 CON G gag. OPT

2005 CON G 349: 021

ATGGGCGCCTGGGCGGCAAGCTGGACGCTGGAGAAAGATCGGCTGGGCCCGGGGCAAGAAAGTACCGCAT  
GAAACACCTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGACCTGCTGGAGACCGCGAGGGCTGCCAGCAGATCATGG  
GCCAGTGCAGCCCGCTTCAGACCGGACCGAGGAGCTCGCTCCCTGTTCAACACCGTGGCCACCTGTACTGCGTGCAACAGCGCATC  
GAGTGAAGGACCAAGGAGGCTGGAGGAGTGGAGAAGATCCAGAAGATCCAGAGAAAGACCCAGCAGGCCGCCATGGACGAGGG  
CAACTCTCCAGGTGCCAGAACTACCCCATCGTGCAAGACGCCACGAGGCCAGATGTTCTCCGCCCTGCCAGGGCCATCTCCCCCGCACCTGAACG  
CCTGGTGAAGGTGGTGGAGGAGAGGCTTCTCCCCGAGGTGATCCCATGTTCTCCGCCCTGCCAGGGCCACCCCGCAGGACCTG  
AACACCATGCTGAACACCGTGGGCGGCCACAGGCCGCCATGCAGTGTGAAGGACACCATCAACGAGAGGCCGCCGAGTGGGACCGCAT  
GCACCCCCAGCAGGCCGCCCATCCCCCGGCCAGATCCGCGAGCCCCCGGCTCCGACATCGCCGACCCCTCCACCTGCAGGAGC  
AGATCCGCTGGATGACCTCAACCCCCCATCCCCGTGGCGAGATCTAAGCGCTGGATCATCTGGGCCCTGAACAAGATCGTGCGCATG  
TACTCCCCGTGCCATCTGGACATCCGCCAGGGCCCAAGGAGCCCTTCCGCGACTACGTGGAACCGCTTCTTCAAGACCTTGCGCCGCCGA  
GCAGGCCACCCAGGAGTGAGGGCTGGATGACCGACACCTGCTGGTGAGAACGCCAACCCGACTGCAAGACCATCTCGCGGCCCTGG  
GCCCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCTGGGCGGCCCTCCCAAGGCCCGCTGCTGGCCGAGGCCATGTCC  
CAGGCTCCGGCGCGCGCCCATCATGATGCAGAAGTCCAACCTCAAGGGCCCCCGCCGACCATCAAGTGTCTCAACTGCGGCAAGGA  
GGGCCACCTGGCCCGCACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGCCACCCAGATGAAGACTGCACCGAGC  
GCCAGGCCAACTCTCTGGGCAAGATCTGGCCCTCCAACAAGGGCGCCCCCGCAACTCTTCAGAAACCGCCCCGAGCCACCGCCCCCCC  
GCCAGTCTTCGGCTTCGGCGAGGAGATCGCCCCCTCCCCCAAGCAGGAGCAGAGGAGCTGTACCCCCCTGGCCTCCCTGAAGTC  
CCTGTTGGCTCCGACCCCTAA

Fig. 72A

## 13. 2003 CON H gag. PEP

MGARASVL<sup>5</sup>SGGKLD<sup>1</sup>AW<sup>2</sup>KE<sup>3</sup>IR<sup>4</sup>LR<sup>5</sup>PG<sup>6</sup>GK<sup>7</sup>KY<sup>8</sup>RL<sup>9</sup>KL<sup>10</sup>VW<sup>11</sup>AS<sup>12</sup>RE<sup>13</sup>LE<sup>14</sup>RF<sup>15</sup>AL<sup>16</sup>NP<sup>17</sup>GL<sup>18</sup>LE<sup>19</sup>TA<sup>20</sup>EG<sup>21</sup>CL<sup>22</sup>QI<sup>23</sup>IE<sup>24</sup>QL<sup>25</sup>PA<sup>26</sup>IK<sup>27</sup>TG<sup>28</sup>TE<sup>29</sup>EL<sup>30</sup>QSL<sup>31</sup>FN<sup>32</sup>TV<sup>33</sup>AV<sup>34</sup>LY<sup>35</sup>CV<sup>36</sup>HQ<sup>37</sup>RI<sup>38</sup>  
 DV<sup>39</sup>KDT<sup>40</sup>KE<sup>41</sup>AL<sup>42</sup>GK<sup>43</sup>IE<sup>44</sup>EQ<sup>45</sup>NS<sup>46</sup>QK<sup>47</sup>TQ<sup>48</sup>AA<sup>49</sup>AD<sup>50</sup>KE<sup>51</sup>KN<sup>52</sup>VS<sup>53</sup>QNP<sup>54</sup>IV<sup>55</sup>QA<sup>56</sup>QGM<sup>57</sup>VHQ<sup>58</sup>AI<sup>59</sup>SP<sup>60</sup>TL<sup>61</sup>NA<sup>62</sup>W<sup>63</sup>K<sup>64</sup>VE<sup>65</sup>EKA<sup>66</sup>FS<sup>67</sup>PE<sup>68</sup>VI<sup>69</sup>PM<sup>70</sup>FS<sup>71</sup>AL<sup>72</sup>SE<sup>73</sup>GAT<sup>74</sup>PQ<sup>75</sup>DL<sup>76</sup>  
 NAM<sup>77</sup>LN<sup>78</sup>TV<sup>79</sup>GG<sup>80</sup>HQA<sup>81</sup>AM<sup>82</sup>QML<sup>83</sup>KD<sup>84</sup>TINE<sup>85</sup>EAE<sup>86</sup>FW<sup>87</sup>DR<sup>88</sup>LHP<sup>89</sup>VHAG<sup>90</sup>PI<sup>91</sup>PP<sup>92</sup>QM<sup>93</sup>RE<sup>94</sup>PR<sup>95</sup>GS<sup>96</sup>DI<sup>97</sup>AG<sup>98</sup>TT<sup>99</sup>SL<sup>100</sup>QE<sup>101</sup>IA<sup>102</sup>W<sup>103</sup>MT<sup>104</sup>GN<sup>105</sup>PP<sup>106</sup>IP<sup>107</sup>VG<sup>108</sup>DI<sup>109</sup>YK<sup>110</sup>RW<sup>111</sup>IL<sup>112</sup>GL<sup>113</sup>NK<sup>114</sup>IV<sup>115</sup>RM<sup>116</sup>  
 YSP<sup>117</sup>VS<sup>118</sup>IL<sup>119</sup>DI<sup>120</sup>KQ<sup>121</sup>GP<sup>122</sup>KE<sup>123</sup>FR<sup>124</sup>DY<sup>125</sup>DR<sup>126</sup>FF<sup>127</sup>KT<sup>128</sup>LR<sup>129</sup>AE<sup>130</sup>QAT<sup>131</sup>QDV<sup>132</sup>KN<sup>133</sup>WT<sup>134</sup>DT<sup>135</sup>LL<sup>136</sup>VQ<sup>137</sup>NA<sup>138</sup>NP<sup>139</sup>DK<sup>140</sup>KT<sup>141</sup>IL<sup>142</sup>RAL<sup>143</sup>GQ<sup>144</sup>AS<sup>145</sup>IE<sup>146</sup>EM<sup>147</sup>MT<sup>148</sup>AC<sup>149</sup>QGV<sup>150</sup>GP<sup>151</sup>SH<sup>152</sup>KAR<sup>153</sup>VLA<sup>154</sup>EAM<sup>155</sup>S  
 QVT<sup>156</sup>NA<sup>157</sup>AA<sup>158</sup>IM<sup>159</sup>MQ<sup>160</sup>GN<sup>161</sup>FK<sup>162</sup>GP<sup>163</sup>RI<sup>164</sup>VK<sup>165</sup>CF<sup>166</sup>NC<sup>167</sup>GK<sup>168</sup>EG<sup>169</sup>HI<sup>170</sup>ARN<sup>171</sup>CR<sup>172</sup>AP<sup>173</sup>RK<sup>174</sup>KG<sup>175</sup>CW<sup>176</sup>KC<sup>177</sup>GR<sup>178</sup>EH<sup>179</sup>QMK<sup>180</sup>DC<sup>181</sup>TER<sup>182</sup>QAN<sup>183</sup>FL<sup>184</sup>GKI<sup>185</sup>WP<sup>186</sup>SS<sup>187</sup>KGR<sup>188</sup>PN<sup>189</sup>FL<sup>190</sup>QSR<sup>191</sup>PE<sup>192</sup>PT<sup>193</sup>APP<sup>194</sup>  
 AES<sup>195</sup>FG<sup>196</sup>GE<sup>197</sup>EM<sup>198</sup>TP<sup>199</sup>SP<sup>200</sup>KQ<sup>201</sup>EL<sup>202</sup>KD<sup>203</sup>KE<sup>204</sup>PP<sup>205</sup>LA<sup>206</sup>SL<sup>207</sup>RS<sup>208</sup>LF<sup>209</sup>GN<sup>210</sup>DL<sup>211</sup>PL<sup>212</sup>SL<sup>213</sup>Q\$

Fig. 72B

## 2003 CON H gag. OPT

AT<sup>1</sup>GG<sup>2</sup>CG<sup>3</sup>C<sup>4</sup>CG<sup>5</sup>GC<sup>6</sup>CT<sup>7</sup>CC<sup>8</sup>GT<sup>9</sup>CT<sup>10</sup>CG<sup>11</sup>CG<sup>12</sup>CG<sup>13</sup>CA<sup>14</sup>AG<sup>15</sup>CT<sup>16</sup>GG<sup>17</sup>AG<sup>18</sup>AG<sup>19</sup>AT<sup>20</sup>CG<sup>21</sup>CC<sup>22</sup>CG<sup>23</sup>CG<sup>24</sup>CG<sup>25</sup>CA<sup>26</sup>AG<sup>27</sup>AAG<sup>28</sup>AT<sup>29</sup>AC<sup>30</sup>CG<sup>31</sup>CC<sup>32</sup>  
 GA<sup>33</sup>AG<sup>34</sup>CA<sup>35</sup>CT<sup>36</sup>GT<sup>37</sup>GG<sup>38</sup>GC<sup>39</sup>CT<sup>40</sup>CC<sup>41</sup>CG<sup>42</sup>AG<sup>43</sup>CT<sup>44</sup>GG<sup>45</sup>AG<sup>46</sup>CT<sup>47</sup>TC<sup>48</sup>GC<sup>49</sup>CT<sup>50</sup>GA<sup>51</sup>AC<sup>52</sup>CC<sup>53</sup>GG<sup>54</sup>CT<sup>55</sup>GT<sup>56</sup>GG<sup>57</sup>AG<sup>58</sup>AC<sup>59</sup>CC<sup>60</sup>CG<sup>61</sup>AG<sup>62</sup>GG<sup>63</sup>CT<sup>64</sup>GC<sup>65</sup>AT<sup>66</sup>CG<sup>67</sup>  
 AG<sup>68</sup>CA<sup>69</sup>GT<sup>70</sup>GC<sup>71</sup>AG<sup>72</sup>CC<sup>73</sup>CG<sup>74</sup>CA<sup>75</sup>AG<sup>76</sup>AC<sup>77</sup>CG<sup>78</sup>CA<sup>79</sup>CG<sup>80</sup>AG<sup>81</sup>AG<sup>82</sup>CT<sup>83</sup>GC<sup>84</sup>AG<sup>85</sup>TC<sup>86</sup>CT<sup>87</sup>TT<sup>88</sup>CA<sup>89</sup>AC<sup>90</sup>CC<sup>91</sup>GT<sup>92</sup>GG<sup>93</sup>CG<sup>94</sup>GT<sup>95</sup>GT<sup>96</sup>ACT<sup>97</sup>GC<sup>98</sup>AT<sup>99</sup>GC<sup>100</sup>CG<sup>101</sup>CA<sup>102</sup>CG<sup>103</sup>CA<sup>104</sup>CG<sup>105</sup>CA<sup>106</sup>CG<sup>107</sup>CA<sup>108</sup>CG<sup>109</sup>CA<sup>110</sup>CG<sup>111</sup>CA<sup>112</sup>CG<sup>113</sup>CA<sup>114</sup>CG<sup>115</sup>CA<sup>116</sup>CG<sup>117</sup>CA<sup>118</sup>CG<sup>119</sup>CA<sup>120</sup>CG<sup>121</sup>CA<sup>122</sup>CG<sup>123</sup>CA<sup>124</sup>CG<sup>125</sup>CA<sup>126</sup>CG<sup>127</sup>CA<sup>128</sup>CG<sup>129</sup>CA<sup>130</sup>CG<sup>131</sup>CA<sup>132</sup>CG<sup>133</sup>CA<sup>134</sup>CG<sup>135</sup>CA<sup>136</sup>CG<sup>137</sup>CA<sup>138</sup>CG<sup>139</sup>CA<sup>140</sup>CG<sup>141</sup>CA<sup>142</sup>CG<sup>143</sup>CA<sup>144</sup>CG<sup>145</sup>CA<sup>146</sup>CG<sup>147</sup>CA<sup>148</sup>CG<sup>149</sup>CA<sup>150</sup>CG<sup>151</sup>CA<sup>152</sup>CG<sup>153</sup>CA<sup>154</sup>CG<sup>155</sup>CA<sup>156</sup>CG<sup>157</sup>CA<sup>158</sup>CG<sup>159</sup>CA<sup>160</sup>CG<sup>161</sup>CA<sup>162</sup>CG<sup>163</sup>CA<sup>164</sup>CG<sup>165</sup>CA<sup>166</sup>CG<sup>167</sup>CA<sup>168</sup>CG<sup>169</sup>CA<sup>170</sup>CG<sup>171</sup>CA<sup>172</sup>CG<sup>173</sup>CA<sup>174</sup>CG<sup>175</sup>CA<sup>176</sup>CG<sup>177</sup>CA<sup>178</sup>CG<sup>179</sup>CA<sup>180</sup>CG<sup>181</sup>CA<sup>182</sup>CG<sup>183</sup>CA<sup>184</sup>CG<sup>185</sup>CA<sup>186</sup>CG<sup>187</sup>CA<sup>188</sup>CG<sup>189</sup>CA<sup>190</sup>CG<sup>191</sup>CA<sup>192</sup>CG<sup>193</sup>CA<sup>194</sup>CG<sup>195</sup>CA<sup>196</sup>CG<sup>197</sup>CA<sup>198</sup>CG<sup>199</sup>CA<sup>200</sup>CG<sup>201</sup>CA<sup>202</sup>CG<sup>203</sup>CA<sup>204</sup>CG<sup>205</sup>CA<sup>206</sup>CG<sup>207</sup>CA<sup>208</sup>CG<sup>209</sup>CA<sup>210</sup>CG<sup>211</sup>CA<sup>212</sup>CG<sup>213</sup>CA<sup>214</sup>CG<sup>215</sup>CA<sup>216</sup>CG<sup>217</sup>CA<sup>218</sup>CG<sup>219</sup>CA<sup>220</sup>CG<sup>221</sup>CA<sup>222</sup>CG<sup>223</sup>CA<sup>224</sup>CG<sup>225</sup>CA<sup>226</sup>CG<sup>227</sup>CA<sup>228</sup>CG<sup>229</sup>CA<sup>230</sup>CG<sup>231</sup>CA<sup>232</sup>CG<sup>233</sup>CA<sup>234</sup>CG<sup>235</sup>CA<sup>236</sup>CG<sup>237</sup>CA<sup>238</sup>CG<sup>239</sup>CA<sup>240</sup>CG<sup>241</sup>CA<sup>242</sup>CG<sup>243</sup>CA<sup>244</sup>CG<sup>245</sup>CA<sup>246</sup>CG<sup>247</sup>CA<sup>248</sup>CG<sup>249</sup>CA<sup>250</sup>CG<sup>251</sup>CA<sup>252</sup>CG<sup>253</sup>CA<sup>254</sup>CG<sup>255</sup>CA<sup>256</sup>CG<sup>257</sup>CA<sup>258</sup>CG<sup>259</sup>CA<sup>260</sup>CG<sup>261</sup>CA<sup>262</sup>CG<sup>263</sup>CA<sup>264</sup>CG<sup>265</sup>CA<sup>266</sup>CG<sup>267</sup>CA<sup>268</sup>CG<sup>269</sup>CA<sup>270</sup>CG<sup>271</sup>CA<sup>272</sup>CG<sup>273</sup>CA<sup>274</sup>CG<sup>275</sup>CA<sup>276</sup>CG<sup>277</sup>CA<sup>278</sup>CG<sup>279</sup>CA<sup>280</sup>CG<sup>281</sup>CA<sup>282</sup>CG<sup>283</sup>CA<sup>284</sup>CG<sup>285</sup>CA<sup>286</sup>CG<sup>287</sup>CA<sup>288</sup>CG<sup>289</sup>CA<sup>290</sup>CG<sup>291</sup>CA<sup>292</sup>CG<sup>293</sup>CA<sup>294</sup>CG<sup>295</sup>CA<sup>296</sup>CG<sup>297</sup>CA<sup>298</sup>CG<sup>299</sup>CA<sup>300</sup>CG<sup>301</sup>CA<sup>302</sup>CG<sup>303</sup>CA<sup>304</sup>CG<sup>305</sup>CA<sup>306</sup>CG<sup>307</sup>CA<sup>308</sup>CG<sup>309</sup>CA<sup>310</sup>CG<sup>311</sup>CA<sup>312</sup>CG<sup>313</sup>CA<sup>314</sup>CG<sup>315</sup>CA<sup>316</sup>CG<sup>317</sup>CA<sup>318</sup>CG<sup>319</sup>CA<sup>320</sup>CG<sup>321</sup>CA<sup>322</sup>CG<sup>323</sup>CA<sup>324</sup>CG<sup>325</sup>CA<sup>326</sup>CG<sup>327</sup>CA<sup>328</sup>CG<sup>329</sup>CA<sup>330</sup>CG<sup>331</sup>CA<sup>332</sup>CG<sup>333</sup>CA<sup>334</sup>CG<sup>335</sup>CA<sup>336</sup>CG<sup>337</sup>CA<sup>338</sup>CG<sup>339</sup>CA<sup>340</sup>CG<sup>341</sup>CA<sup>342</sup>CG<sup>343</sup>CA<sup>344</sup>CG<sup>345</sup>CA<sup>346</sup>CG<sup>347</sup>CA<sup>348</sup>CG<sup>349</sup>CA<sup>350</sup>CG<sup>351</sup>CA<sup>352</sup>CG<sup>353</sup>CA<sup>354</sup>CG<sup>355</sup>CA<sup>356</sup>CG<sup>357</sup>CA<sup>358</sup>CG<sup>359</sup>CA<sup>360</sup>CG<sup>361</sup>CA<sup>362</sup>CG<sup>363</sup>CA<sup>364</sup>CG<sup>365</sup>CA<sup>366</sup>CG<sup>367</sup>CA<sup>368</sup>CG<sup>369</sup>CA<sup>370</sup>CG<sup>371</sup>CA<sup>372</sup>CG<sup>373</sup>CA<sup>374</sup>CG<sup>375</sup>CA<sup>376</sup>CG<sup>377</sup>CA<sup>378</sup>CG<sup>379</sup>CA<sup>380</sup>CG<sup>381</sup>CA<sup>382</sup>CG<sup>383</sup>CA<sup>384</sup>CG<sup>385</sup>CA<sup>386</sup>CG<sup>387</sup>CA<sup>388</sup>CG<sup>389</sup>CA<sup>390</sup>CG<sup>391</sup>CA<sup>392</sup>CG<sup>393</sup>CA<sup>394</sup>CG<sup>395</sup>CA<sup>396</sup>CG<sup>397</sup>CA<sup>398</sup>CG<sup>399</sup>CA<sup>400</sup>CG<sup>401</sup>CA<sup>402</sup>CG<sup>403</sup>CA<sup>404</sup>CG<sup>405</sup>CA<sup>406</sup>CG<sup>407</sup>CA<sup>408</sup>CG<sup>409</sup>CA<sup>410</sup>CG<sup>411</sup>CA<sup>412</sup>CG<sup>413</sup>CA<sup>414</sup>CG<sup>415</sup>CA<sup>416</sup>CG<sup>417</sup>CA<sup>418</sup>CG<sup>419</sup>CA<sup>420</sup>CG<sup>421</sup>CA<sup>422</sup>CG<sup>423</sup>CA<sup>424</sup>CG<sup>425</sup>CA<sup>426</sup>CG<sup>427</sup>CA<sup>428</sup>CG<sup>429</sup>CA<sup>430</sup>CG<sup>431</sup>CA<sup>432</sup>CG<sup>433</sup>CA<sup>434</sup>CG<sup>435</sup>CA<sup>436</sup>CG<sup>437</sup>CA<sup>438</sup>CG<sup>439</sup>CA<sup>440</sup>CG<sup>441</sup>CA<sup>442</sup>CG<sup>443</sup>CA<sup>444</sup>CG<sup>445</sup>CA<sup>446</sup>CG<sup>447</sup>CA<sup>448</sup>CG<sup>449</sup>CA<sup>450</sup>CG<sup>451</sup>CA<sup>452</sup>CG<sup>453</sup>CA<sup>454</sup>CG<sup>455</sup>CA<sup>456</sup>CG<sup>457</sup>CA<sup>458</sup>CG<sup>459</sup>CA<sup>460</sup>CG<sup>461</sup>CA<sup>462</sup>CG<sup>463</sup>CA<sup>464</sup>CG<sup>465</sup>CA<sup>466</sup>CG<sup>467</sup>CA<sup>468</sup>CG<sup>469</sup>CA<sup>470</sup>CG<sup>471</sup>CA<sup>472</sup>CG<sup>473</sup>CA<sup>474</sup>CG<sup>475</sup>CA<sup>476</sup>CG<sup>477</sup>CA<sup>478</sup>CG<sup>479</sup>CA<sup>480</sup>CG<sup>481</sup>CA<sup>482</sup>CG<sup>483</sup>CA<sup>484</sup>CG<sup>485</sup>CA<sup>486</sup>CG<sup>487</sup>CA<sup>488</sup>CG<sup>489</sup>CA<sup>490</sup>CG<sup>491</sup>CA<sup>492</sup>CG<sup>493</sup>CA<sup>494</sup>CG<sup>495</sup>CA<sup>496</sup>CG<sup>497</sup>CA<sup>498</sup>CG<sup>499</sup>CA<sup>500</sup>CG<sup>501</sup>CA<sup>502</sup>CG<sup>503</sup>CA<sup>504</sup>CG<sup>505</sup>CA<sup>506</sup>CG<sup>507</sup>CA<sup>508</sup>CG<sup>509</sup>CA<sup>510</sup>CG<sup>511</sup>CA<sup>512</sup>CG<sup>513</sup>CA<sup>514</sup>CG<sup>515</sup>CA<sup>516</sup>CG<sup>517</sup>CA<sup>518</sup>CG<sup>519</sup>CA<sup>520</sup>CG<sup>521</sup>CA<sup>522</sup>CG<sup>523</sup>CA<sup>524</sup>CG<sup>525</sup>CA<sup>526</sup>CG<sup>527</sup>CA<sup>528</sup>CG<sup>529</sup>CA<sup>530</sup>CG<sup>531</sup>CA<sup>532</sup>CG<sup>533</sup>CA<sup>534</sup>CG<sup>535</sup>CA<sup>536</sup>CG<sup>537</sup>CA<sup>538</sup>CG<sup>539</sup>CA<sup>540</sup>CG<sup>541</sup>CA<sup>542</sup>CG<sup>543</sup>CA<sup>544</sup>CG<sup>545</sup>CA<sup>546</sup>CG<sup>547</sup>CA<sup>548</sup>CG<sup>549</sup>CA<sup>550</sup>CG<sup>551</sup>CA<sup>552</sup>CG<sup>553</sup>CA<sup>554</sup>CG<sup>555</sup>CA<sup>556</sup>CG<sup>557</sup>CA<sup>558</sup>CG<sup>559</sup>CA<sup>560</sup>CG<sup>561</sup>CA<sup>562</sup>CG<sup>563</sup>CA<sup>564</sup>CG<sup>565</sup>CA<sup>566</sup>CG<sup>567</sup>CA<sup>568</sup>CG<sup>569</sup>CA<sup>570</sup>CG<sup>571</sup>CA<sup>572</sup>CG<sup>573</sup>CA<sup>574</sup>CG<sup>575</sup>CA<sup>576</sup>CG<sup>577</sup>CA<sup>578</sup>CG<sup>579</sup>CA<sup>580</sup>CG<sup>581</sup>CA<sup>582</sup>CG<sup>583</sup>CA<sup>584</sup>CG<sup>585</sup>CA<sup>586</sup>CG<sup>587</sup>CA<sup>588</sup>CG<sup>589</sup>CA<sup>590</sup>CG<sup>591</sup>CA<sup>592</sup>CG<sup>593</sup>CA<sup>594</sup>CG<sup>595</sup>CA<sup>596</sup>CG<sup>597</sup>CA<sup>598</sup>CG<sup>599</sup>CA<sup>600</sup>CG<sup>601</sup>CA<sup>602</sup>CG<sup>603</sup>CA<sup>604</sup>CG<sup>605</sup>CA<sup>606</sup>CG<sup>607</sup>CA<sup>608</sup>CG<sup>609</sup>CA<sup>610</sup>CG<sup>611</sup>CA<sup>612</sup>CG<sup>613</sup>CA<sup>614</sup>CG<sup>615</sup>CA<sup>616</sup>CG<sup>617</sup>CA<sup>618</sup>CG<sup>619</sup>CA<sup>620</sup>CG<sup>621</sup>CA<sup>622</sup>CG<sup>623</sup>CA<sup>624</sup>CG<sup>625</sup>CA<sup>626</sup>CG<sup>627</sup>CA<sup>628</sup>CG<sup>629</sup>CA<sup>630</sup>CG<sup>631</sup>CA<sup>632</sup>CG<sup>633</sup>CA<sup>634</sup>CG<sup>635</sup>CA<sup>636</sup>CG<sup>637</sup>CA<sup>638</sup>CG<sup>639</sup>CA<sup>640</sup>CG<sup>641</sup>CA<sup>642</sup>CG<sup>643</sup>CA<sup>644</sup>CG<sup>645</sup>CA<sup>646</sup>CG<sup>647</sup>CA<sup>648</sup>CG<sup>649</sup>CA<sup>650</sup>CG<sup>651</sup>CA<sup>652</sup>CG<sup>653</sup>CA<sup>654</sup>CG<sup>655</sup>CA<sup>656</sup>CG<sup>657</sup>CA<sup>658</sup>CG<sup>659</sup>CA<sup>660</sup>CG<sup>661</sup>CA<sup>662</sup>CG<sup>663</sup>CA<sup>664</sup>CG<sup>665</sup>CA<sup>666</sup>CG<sup>667</sup>CA<sup>668</sup>CG<sup>669</sup>CA<sup>670</sup>CG<sup>671</sup>CA<sup>672</sup>CG<sup>673</sup>CA<sup>674</sup>CG<sup>675</sup>CA<sup>676</sup>CG<sup>677</sup>CA<sup>678</sup>CG<sup>679</sup>CA<sup>680</sup>CG<sup>681</sup>CA<sup>682</sup>CG<sup>683</sup>CA<sup>684</sup>CG<sup>685</sup>CA<sup>686</sup>CG<sup>687</sup>CA<sup>688</sup>CG<sup>689</sup>CA<sup>690</sup>CG<sup>691</sup>CA<sup>692</sup>CG<sup>693</sup>CA<sup>694</sup>CG<sup>695</sup>CA<sup>696</sup>CG<sup>697</sup>CA<sup>698</sup>CG<sup>699</sup>CA<sup>700</sup>CG<sup>701</sup>CA<sup>702</sup>CG<sup>703</sup>CA<sup>704</sup>CG<sup>705</sup>CA<sup>706</sup>CG<sup>707</sup>CA<sup>708</sup>CG<sup>709</sup>CA<sup>710</sup>CG<sup>711</sup>CA<sup>712</sup>CG<sup>713</sup>CA<sup>714</sup>CG<sup>715</sup>CA<sup>716</sup>CG<sup>717</sup>CA<sup>718</sup>CG<sup>719</sup>CA<sup>720</sup>CG<sup>721</sup>CA<sup>722</sup>CG<sup>723</sup>CA<sup>724</sup>CG<sup>725</sup>CA<sup>726</sup>CG<sup>727</sup>CA<sup>728</sup>CG<sup>729</sup>CA<sup>730</sup>CG<sup>731</sup>CA<sup>732</sup>CG<sup>733</sup>CA<sup>734</sup>CG<sup>735</sup>CA<sup>736</sup>CG<sup>737</sup>CA<sup>738</sup>CG<sup>739</sup>CA<sup>740</sup>CG<sup>741</sup>CA<sup>742</sup>CG<sup>743</sup>CA<sup>744</sup>CG<sup>745</sup>CA<sup>746</sup>CG<sup>747</sup>CA<sup>748</sup>CG<sup>749</sup>CA<sup>750</sup>CG<sup>751</sup>CA<sup>752</sup>CG<sup>753</sup>CA<sup>754</sup>CG<sup>755</sup>CA<sup>756</sup>CG<sup>757</sup>CA<sup>758</sup>CG<sup>759</sup>CA<sup>760</sup>CG<sup>761</sup>CA<sup>762</sup>CG<sup>763</sup>CA<sup>764</sup>CG<sup>765</sup>CA<sup>766</sup>CG<sup>767</sup>CA<sup>768</sup>CG<sup>769</sup>CA<sup>770</sup>CG<sup>771</sup>CA<sup>772</sup>CG<sup>773</sup>CA<sup>774</sup>CG<sup>775</sup>CA<sup>776</sup>CG<sup>777</sup>CA<sup>778</sup>CG<sup>779</sup>CA<sup>780</sup>CG<sup>781</sup>CA<sup>782</sup>CG<sup>783</sup>CA<sup>784</sup>CG<sup>785</sup>CA<sup>786</sup>CG<sup>787</sup>CA<sup>788</sup>CG<sup>789</sup>CA<sup>790</sup>CG<sup>791</sup>CA<sup>792</sup>CG<sup>793</sup>CA<sup>794</sup>CG<sup>795</sup>CA<sup>796</sup>CG<sup>797</sup>CA<sup>798</sup>CG<sup>799</sup>CA<sup>800</sup>CG<sup>801</sup>CA<sup>802</sup>CG<sup>803</sup>CA<sup>804</sup>CG<sup>805</sup>CA<sup>806</sup>CG<sup>807</sup>CA<sup>808</sup>CG<sup>809</sup>CA<sup>810</sup>CG<sup>811</sup>CA<sup>812</sup>CG<sup>813</sup>CA<sup>814</sup>CG<sup>815</sup>CA<sup>816</sup>CG<sup>817</sup>CA<sup>818</sup>CG<sup>819</sup>CA<sup>820</sup>CG<sup>821</sup>CA<sup>822</sup>CG<sup>823</sup>CA<sup>824</sup>CG<sup>825</sup>CA<sup>826</sup>CG<sup>827</sup>CA<sup>828</sup>CG<sup>829</sup>CA<sup>830</sup>CG<sup>831</sup>CA<sup>832</sup>CG<sup>833</sup>CA<sup>834</sup>CG<sup>835</sup>CA<sup>836</sup>CG<sup>837</sup>CA<sup>838</sup>CG<sup>839</sup>CA<sup>840</sup>CG<sup>841</sup>CA<sup>842</sup>CG<sup>843</sup>CA<sup>844</sup>CG<sup>845</sup>CA<sup>846</sup>CG<sup>847</sup>CA<sup>848</sup>CG<sup>849</sup>CA<sup>850</sup>CG<sup>851</sup>CA<sup>852</sup>CG<sup>853</sup>CA<sup>854</sup>CG<sup>855</sup>CA<sup>856</sup>CG<sup>857</sup>CA<sup>858</sup>CG<sup>859</sup>CA<sup>860</sup>CG<sup>861</sup>CA<sup>862</sup>CG<sup>863</sup>CA<sup>864</sup>CG<sup>865</sup>CA<sup>866</sup>CG<sup>867</sup>CA<sup>868</sup>CG<sup>869</sup>CA<sup>870</sup>CG<sup>871</sup>CA<sup>872</sup>CG<sup>873</sup>CA<sup>874</sup>CG<sup>875</sup>CA<sup>876</sup>CG<sup>877</sup>CA<sup>878</sup>CG<sup>879</sup>CA<sup>880</sup>CG<sup>881</sup>CA<sup>882</sup>CG<sup>883</sup>CA<sup>884</sup>CG<sup>885</sup>CA<sup>886</sup>CG<sup>887</sup>CA<sup>888</sup>CG<sup>889</sup>CA<sup>890</sup>CG<sup>891</sup>CA<sup>892</sup>CG<sup>893</sup>CA<sup>894</sup>CG<sup>895</sup>CA<sup>896</sup>CG<sup>897</sup>CA<sup>898</sup>CG<sup>899</sup>CA<sup>900</sup>CG<sup>901</sup>CA<sup>902</sup>CG<sup>903</sup>CA<sup>904</sup>CG<sup>905</sup>CA<sup>906</sup>CG<sup>907</sup>CA<sup>908</sup>CG<sup>909</sup>CA<sup>910</sup>CG<sup>911</sup>CA<sup>912</sup>CG<sup>913</sup>CA<sup>914</sup>CG<sup>915</sup>CA<sup>916</sup>CG<sup>917</sup>CA<sup>918</sup>CG<sup>919</sup>CA<sup>920</sup>CG<sup>921</sup>CA<sup>922</sup>CG<sup>923</sup>CA<sup>924</sup>CG<sup>925</sup>CA<sup>926</sup>CG<sup>927</sup>CA<sup>928</sup>CG<sup>929</sup>CA<sup>930</sup>CG<sup>931</sup>CA<sup>932</sup>CG<sup>933</sup>CA<sup>934</sup>CG<sup>935</sup>CA<sup>936</sup>CG<sup>937</sup>CA<sup>938</sup>CG<sup>939</sup>CA<sup>940</sup>CG<sup>941</sup>CA<sup>942</sup>CG<sup>943</sup>CA<sup>944</sup>CG<sup>945</sup>CA<sup>946</sup>CG<sup>947</sup>CA<sup>948</sup>CG<sup>949</sup>CA<sup>950</sup>CG<sup>951</sup>CA<sup>952</sup>CG<sup>953</sup>CA<sup>954</sup>CG<sup>955</sup>CA<sup>956</sup>CG<sup>957</sup>CA<sup>958</sup>CG<sup>959</sup>CA<sup>960</sup>CG<sup>961</sup>CA<sup>962</sup>CG<sup>963</sup>CA<sup>964</sup>CG<sup>965</sup>CA<sup>966</sup>CG<sup>967</sup>CA<sup>968</sup>CG<sup>969</sup>CA<sup>970</sup>CG<sup>971</sup>CA<sup>972</sup>CG<sup>973</sup>CA<sup>974</sup>CG<sup>975</sup>CA<sup>976</sup>CG<sup>977</sup>CA<sup>978</sup>CG<sup>979</sup>CA<sup>980</sup>CG<sup>981</sup>CA<sup>982</sup>CG<sup>983</sup>CA<sup>984</sup>CG<sup>985</sup>CA<sup>986</sup>CG<sup>987</sup>CA<sup>988</sup>CG<sup>989</sup>CA<sup>990</sup>CG<sup>991</sup>CA<sup>992</sup>CG<sup>993</sup>CA<sup>994</sup>CG<sup>995</sup>CA<sup>996</sup>CG<sup>997</sup>CA<sup>998</sup>CG<sup>999</sup>CA<sup>1000</sup>CG<sup>1001</sup>CA<sup>1002</sup>CG<sup>1003</sup>CA<sup>1004</sup>CG<sup>1005</sup>CA<sup>1006</sup>CG<sup>1007</sup>CA<sup>1008</sup>CG<sup>1009</sup>

Fig. 73A

## 14. 2003 CON K gag. PEP

MGARASVLSSGGLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPSSLLETEGCRQIIROQLPSLQTSSEELKSLFNTVATLYCVHQRI  
 EVRDTKEALDKLEEQNKSOQKTQETADKGVSONYPIVONLQGMVHQALSPTLNWVKVIEEKAFSPEVIPMFSALESEGATPDQDINTML  
 NTVGGHQAAQMMLKDTINEEAAEWDRLLHPVHAGPIPPGOMREPRGSDIAGTSTLQEQITWMTSNPPVPVGEIYKRWIILGLNKIVRMYSVP  
 SILDIRQGPKEPRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQANPDKTILKALGPASLEEMMTACQGVGPGHKARILAEAMSQVTN  
 TAVMMQRGNFKGQRKIIKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGPRGNFLQSRPEPTAPPAESFGE  
 GEEITPSPRQETKDEQGPPLTSLKSLFGNDPLSQ\$

Fig. 73B

## 2003 CON K gag. OPT

ATGGGCGCCCGCGCCTCCGTGCTGCCGGCGGAAGCTGGACACCTGGGAGAAGATCCGCCCTGCCGCCCGGGGCAAGAAGTACCGCCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCAACGAGGGCTGCCCGCAGATCATCC  
 GCCAGTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGAAGTCCCTGTCAACACCCGTGGCCACCCCTGTACTGCGTGCAACGAGCGCATC  
 GAGGTGCGCGACACCAAGGAGGCCCTGGACAACTGGAGGAGGAGCAGAACTCCAGCAGAAGACCCAGAGAGACCGCGGACACCGGACAAAGG  
 CGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACACGAGCCCTGTCCCCCGCACCCCTGAACGCCCTGGGTGAAGG  
 TGATCGAGGAGAAAGCCCTTCTCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGGACCGCTGCACCCCGTGCA  
 AACACCGTGGCGGCCACAGGCCGCCATGCAGATGCTGAAGACACCATCAACGAGGAGGCGCGGAGTGGACCGCTGCACCCCGTGCA  
 CGCCGGCCCATCCCCCGGCCAGATGCGCGAGCCCGCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGCAGATCACCTCGTG  
 TGACCTCCAAACCCCGTGCCTGGCGGAGATCTACAAGCGTGGATCATCTGGGCTGAACAAGATCGTGGCATGTACTCCCCCGTG  
 TCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCTTCCGCCCGGAGCAGGCCACCCA  
 GGAGTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAGAACGCCAACCCGACTGCAAGACCATCTGAAGGCCCTGGGCCCGCGCCCT  
 CCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCGCGCACAAAGGCCGCAATCTGCGGAGGCCATGTCCAGGTGACCAAC  
 ACCGCCGTGATGATGACGCGCGCAACTTCAAGGGCCAGCGCAAGATCATCAAGTGTCTCACTGCGGCAAGGAGGCCACATCGCCCCGCAA  
 CTGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGACTGCACCGAGCGCCAGGCCAACTTCTCTGG  
 GCAAGATCTGGCCCTCCAACAAGGGCCCGCCCGCAACTTCTGCAGTCCGCCCCGAGCCACCGCCCCCGCGGAGTCTCTCGGCTTC  
 GCGGAGGAGATCACCCCTCCCCCGCCAGGAGACCAAGGAGCAGGGCCCCCCCCCTGACCTCCCTGAAGTCCCTGTTCGGGCAACGA  
 CCCCCTGTCCCAGTAA

Fig. 74A

15. 2003 CON 01 AE gag.PEP  
 MGARASVL<sup>~</sup>SGGKLD<sup>~</sup>AW<sup>~</sup>EKIRL<sup>~</sup>RPGGKKYRMKHLVWASRELERFALNPGLETAEGCQ<sup>~</sup>QII<sup>~</sup>EQ<sup>~</sup>LQ<sup>~</sup>STLKTGSEELKSLFNTVATLWCVHQRI  
 EVKDTKEALDKIEEVQNK<sup>~</sup>SQKTQ<sup>~</sup>AAAGT<sup>~</sup>SSSKVSQNYPIVQNAQGMVHQPLSPRTLNAWVKVVEEKGFNPEVIPMFSALSEGATPQDL  
 NMMLNIVGHHQAAMQ<sup>~</sup>MLKETINEEAAEWD<sup>~</sup>RVHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIGWMTNNPPIPVGDIYKRWIIILGLNKIVRM  
 YSPVSI<sup>~</sup>LDIRQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETILLVQANANPDCKSILKALGTGATLEEMMTACQGVGGPSHKARVLA<sup>~</sup>EAMS  
 QAQHANIMMQRGNFQ<sup>~</sup>KRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKG<sup>~</sup>RPGNFPQSRPEPTAPPAEN  
 WGMGEEITSLPKQEQKDKEHPPLVLSKSLFGNDPLSQ\$

Fig. 74B

2003 CON 01 AE gag.OPT  
 ATGGGCGC<sup>~</sup>CGG<sup>~</sup>CCCTCCGTGCTCCGGCGGCAAGCTGGACGCTGGGAGAAGATCCGCCCTGCCGCCCGCGGCGCAAGAAGTACCGCAT  
 GAAGCACCTGGTGTGGGCTCCCGGAGCTGGAGGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCGGAGGGCTGCCAGAGATCATCG  
 AGCAGCTGCAGTCCACCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCACACCGTGGCCACCTGTGGTGGTGCACACGCGCATC  
 GAGGTGAAGGACACCAAGGAGGCCCTGGACAAAGATCGAGGAGGTGCAGAA<sup>~</sup>CAAGTCCAGCAGAAGACCCAGCAGGCCCGCGCCGACCCG  
 CTCCTCCTCCAAGGTGTCCAGAACTACCCATCTGTGCAGAAAGCCAGGCCAGATGGTGCACCAAGCCCTGTCCCCCGCACCTGAACG  
 CCTGGGTGAAGGTGGTGGAGGAGAGGGCTTCAACCCCGAGGTGATCCCATGTTCCTCGCCCTGTCCGAGGGCGCACCCCGCAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACAGCGCGCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCCGCGAGTGGGACCGCGT  
 GCACCCCGTGCA<sup>~</sup>CGCCCGCCCATCCCCCGGCGAGATGCGCGAGCCCCCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGC  
 AGATCGGCTGGATGACCAACAACCCCCCATCCCCGTGGGCGACATCTACAAGCGTGGATCATCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCAAGGAGCCCTTCGCGGACTACGTGGACCGCTTCTACAAGACCCCTGCCGCCCGA  
 GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGGAGACCTTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGAAGGCCCTGG  
 GCACCGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCACAAGGCCCGCGTGTGCCGAGGCCATGTCC  
 CAGGCCAGCACGCCAAACATCATGATGCAGCGCGGCAACTCAAGGCCAGAAAGCGCATCAAGTGTCTCAACTGGGCAAGGAGGCCACCT  
 GGCCCGCAACTGCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGGCCACAGATGAAGGACTGCACCGAGCGCCAGGCCA  
 ACTTCTGGGCAAGATCTGGCCCTCCAAACAAGGGCGCCCCGGCAACTTCCCCAGTCCCCCGGAGCCACCGCCCCCGCGGAGAAC  
 TGGGGCATGGGCGAGGAGATCACCTCCCTGCCCAAGCAGGAGCAGAAGGACACCCCCCCCCCTGGTGTCCCTGAAGTCCCTGTT  
 CGGCAACGACCCCCCTGTCCCAAGTAA

117/178

Fig. 75A

16. 2003 CON 02 AG gag . PEP  
 MGRASVLGGKLD~~AW~~EKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETAEGCQIQMEQLQSALRTGSEELKSLYNTVATLWCVHORI  
 DIKDTKEALDKIEEVONKSKQKTQAAAAATGSSSQNYPIVQNAQGMTHQSMSPRTLNAWKVIEEKAFSPEVIPMFSALESEGATPQDLNMM  
 LNI VGGHQAAMQMLKDTINEEAAEWDVRVHPVHAGPIPPGQMRPRGSDIAGTTSTLQEIQIGWMTSNPPIPVGEIYKRWIVLGNKIVRMYS  
 VSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQNANPDCKSILRALPGATLEEMTACQGVGGPGHKARVLAEAMSQVQ  
 QSNIMQRGNFRGQRTIKFCNCGKEGHLARNCKAPRKKGWCCKGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFPQSRPEPTAPAESFGM  
 GEEITSSPKQEPDRDKGLYPPLTSLKSLFGNDP\$

Fig. 75B

2003 CON 02 AG gag . OPT  
 ATGGCGCCCGCCCTCCGTGCTGCCGGCGCAAGCTGGACGCCCTGGGAGAAGATCCGCCCTGCCGCCCGGGCGGCAAGAAGTACCGCCT  
 GAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG  
 AGCAGCTGAGTCCGCCCTCGCACCCGGCTCCGAGGAGCTGAGTCCCTGTACAACACCGTGGCCACCCCTGTGGTGGCTGCCACGCGCATC  
 GACATCAAGGACACCAAGAGGCCCTGGACAAGATCGAGGAGTGCAGAACAAAGTCCAAGCAGAAGACCCAGAGGCCGCCGCCGCCACCGG  
 CTCCTCTCCAGAACTACCCCATCGTGCAGAACGCCAGGCCAGATGACCCACAGTCCATGTCCCCCGCACCCCTGAACGCCCTGGGTGA  
 AGTGATCGAGGAGAAGGCTTCTCCCCGAGGTGATCCCCATGTCTCCGCCCTGTCCGAGGGCGCCACCCCCAGGACCTGAACATGATG  
 CTGAACATCGTGGCGGCCACCAAGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGGTGCACCCCGT  
 GCACGCCGCCCATCCCCCGGCCAGATGCGGAGCCCCGGCTCCGACATCGCCGGCACCATCCACCTGCAAGGAGCAGATCGGCT  
 GGATGACCTCCAACCCCATCCCGTGGCGAGATCTACAAGCGCTGGATCGTGGCCCTGAACAAGATCGTGCATGTACTCCCC  
 GTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCGCGGACTACGTGGACCGCTTCTTCAAGACCCCTGCGCGCCGAGAGGCCAC  
 CCAGGAGGTGAGAACTGGATACCGAGACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGGCCCTGGGCCCGCGG  
 CCACCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCGCCACCAAGGCCCGCGTGTGGCCGAGGCCATGTCCAGGTGCAG  
 CAGTCCAACATCATGATGCAGCGCGGCAACTCCCGGCCAGCGCACCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACCTGGCCCCGCAA  
 CTGCAAGGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCAAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCCTGG  
 GCAAGATCTGGCCCTCTCCAGGGCGGCCCGGCAACTTCCCCAGTCCCGCCGAGCCACCGGCCCGCCGAGTCTTCGGCATG  
 GCGGAGGAGATCACCTCTCCCCCAAGCAGGAGCCCCCGGACAAAGGGCCTGTACCCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCTAA

118/178

Fig. 76A

## 17. 2003 CON 03 ABG gag .PEP

MGARASVL<sup>1</sup>SGGKLD<sup>2</sup>AW<sup>3</sup>EKIRLRPGGKKYRIKHLVWASRELERFALNP<sup>4</sup>SLLETSEGCQ<sup>5</sup>QILEQLQ<sup>6</sup>PTLKTGSEELKSLYNTVATLYCVHQRI  
 EIKDTKEALDKIEEIQNKSKQKTQAAATGTGSSSKVSQNYPIVQNAQGMTHQSMSPRTLN<sup>7</sup>AWKVIEEKA<sup>8</sup>FSPEVIPMF<sup>9</sup>SALSEGATPQDL  
 NMMLNIVGGHQ<sup>10</sup>AAQM<sup>11</sup>LKDTINEEAAEW<sup>12</sup>DR<sup>13</sup>LHPAQAGFP<sup>14</sup>PGQMREPRGSDIAGTT<sup>15</sup>STLQEQIGWMTSNPPIPVGDIYKRWIILGLNKIVRM  
 YSPVSI<sup>16</sup>LDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTETLLVQ<sup>17</sup>ANPDCKTILRALGSGATLEEMMTACQGVGPGHKARVLAEAMS  
 QVQ<sup>18</sup>ANIMMQKSNFRGP<sup>19</sup>KRIKCFNCGKDGHLARNCRAPRKKGKWC<sup>20</sup>KGKEGHQMKDCTERQANFLGRINPSSKGRPNFPQSRPEPSAPPAEN  
 FGMGEIITPSLKQEQKDREQHPPSISLKS<sup>21</sup>LFNDPLSQ\$

Fig. 76B

## 2003 CON 03 ABG gag .OPT

ATGGGCGC<sup>1</sup>CGCGCCTCCGTGCTGTCGGGGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGTACCGCAT  
 CAAGCACCTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCTCCGAGGGCTGCCAGCAGATCCTGG  
 AGCAGCTGACGCCACCCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTACAACACCGTGGCCACCCCTGTACTGCGTGCAACCGCATC  
 GAGATCAAGGACACCAAGGAGCCCTGGACAAGATCGAGGAGATCCAGAACAAAGTCCAAGCAGAAGACCCAGCAGGCGCCACCCGGCACCCG  
 CTCTCCTCCAAAGGTGTCCAGAACTACCCCATCGTGCAGAACGCCAGGCCAGATGACCCACCAAGTCCATGTCCCCCGGCAACCCCTGAACG  
 CCTGGGTGAAGTGATCGAGGAGAAGGCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGCGGCCACCAAGCGCCCATGCAAGATGCTGAAGGACACCATCAACGAGGAGGCGCGGAGTGGGACCGCTG  
 GCACCCCGCCAGGCGGCCCCCTTCCCCCCCCCATCCCGTGGCGGACATCTACAGCGCTGGATCATCTTGGGCTGAACAAGATCGTGCGCATG  
 AGATCGGCTGGATGACCTCCAACCCCTGACATCCGCCAGGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTCAAGACCTGCGGCGCGA  
 TACTCCCCGTGTCCATCCTGGACATCCGCCAGGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTCAAGACCTGCGGCGCGA  
 GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCTTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGGCGCCCTGG  
 GCTCCGGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGCGTGGCGGCGGCCCGGCCCAAGGCCCGCTGTGCGGAGGCCCATGTCC  
 CAGGTGCAGAACGCCCAACATCATGATGCAGAAAGTCCAACCTCCGCGGCCCAAGCGCATCAAGTGTCTCACTGCGGCAAGGACGGCCACCT  
 GGGCCGCAACTGCCGCGCATCTGGCCCTCCCTCCAAGGGCGGCCCGGCAACTTCCCCAGTCCCCCGGAGCCCTCCGCCCCCGGAGAAC  
 ACTTCC<sup>22</sup>TGGGCGCATCTGGCCCTCCCTCCAAGGGCGGCCCGGCAACTTCCCCAGTCCCCCGGAGCCCTCCGCCCCCGGAGAAC  
 TTCGGCATGGCGGAGGAGATCACCCCTCCCTGAAGCAGGAGCAGAGGACCCCGGAGCACCCCTCCATCTCCCTGAAGTCCCTGT  
 CGGCAACGACCCCTGTCCAGTAA

119/178



Fig. 77A

18. 2003 CON 04 CFX gag. PEP  
 MGARASVLSGGKLDÄWERIRLRPGGKKRYRLKHLVWASRELERFALNPGLLETAEGCQQLMEQLQSTLKTGSEELKSLFNITATLWCVHQRI  
 DVKDTKEALDKVEEMQNKSQKTKQAAADTGGSSNVSONYPIVQNAQGMVHOSISPRTLNAWVKVIEEKAFSPEVIPMFSAISEGATPQDL  
 NMMLNIVGGHQAAMQLKDTINEEAAEWDRAPVHAGPIPPGQMPREPRGSDIAGTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSILDIRQGPKEPFRDYVDRFFKCLRAEQATQEVKNWMTETLLVQANANPDCKSILKALGTGATLEEMMTACQGVGGPSHKARVLAEMS  
 QASNAAAAIMMQSNFKGQRRRIKCFNCGKEGHLARNCRAPRKKGCWKCKEGHQMKDCTERQANFLGRMWPPSSKGRPGNFGFLQSRPEPTAPP  
 AESLEMKETTSPPKQEPDRKELYPLTSLKSLFGSDPLSQS

Fig. 77B

2003 CON 04 CFX gag. OPT  
 ATGGGCGCGCGCTCCGTGCTGTCCGGGGCAAGCTGGACGCTGGGAGCGCATCCGCCCTGCGCCCCGGCGGCAAGAAGTACCGCCT  
 GAAGCACCTGGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCCGAGGGCTGCCAGCAGCTGATGG  
 AGCAGTGCAGTCCACCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCATCGCCACCTGTGGTGGTGCACCCAGCGCATC  
 GACGTGAAGGACACCAAGGAGGCTTGGACAAGTGGAGGAGATGCAGAACAAAGTCCAAAGCAGAACCCAGCAGGCGCGCCGACACCCGG  
 CGGCTCCTCAACGTGTCCAGAACTACCCCATCGTGCAGAACCGCCAGGCGCAGATGGTGCACCATCTCCCGCCGACCCCTGAAACG  
 CCTGGGTGAAGTGTATCGAGGAGAAGCCCTTCTCCCGGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGGACCTG  
 AACATGATGCTGAACATCGTGGGCGGCCACCAAGCCGCTATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGGACCGCGC  
 CCAACCGGTGCACGCGGCCCATCCCCCGGCCAGATGCGCGAGCCCCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGC  
 AGATCGGTGGATGACCTCCAAACCCCATCCCCGTGGCGGAGATCTACAAGCGTGGATCATCTGGGCTGAACAAGATCGTGGCGCATG  
 TACTCCCCCGTGTCCATCTGGACATCCGCCAGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGTGCCTGCGCGCCGA  
 GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGGAGACCTGCTGGTGCAGAACGCCAACCCGACTGCAAGTCCATCCTGAAGGCCCTGG  
 GCACCGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGCGTGGGCGGCCCTCCCAAGGCCCGCGTGTGGCGGAGGCCATGTCC  
 CAGGCTCCAAAGCGCGCCGCTCATGATGCAGAACTCAACTCAAGGGCCAGCGCCGATCATCAAGTGTCTCAACTGCGGCAAGGA  
 GGGCCACCTGGCCGCAACTGCCGCGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCCAGATGAAGGACTGCACCGAGC  
 GCCAGGCCAACTTCTGGGCGCATGTGGCCCTCTCTCAAGGGCGCCCCGGCAACTTCTGTGAGTCCCGCCCCGAGCCACCGCCCCCCC  
 GCCGAGTCCCTGGAGATGAAGGAGGAGACCACTCTCTCCCCCAAGCAGGAGCCCCCGGACAAGGAGCTGTACCCCTGACCTCCCTGAAGTC  
 CCTGTTGGGCTCCGACCCCTGTCCCAGTAA

120/178

Fig. 78A

## 19. 2003\_CON\_06\_CPX\_gag.gag.PEP

MGARASVL<sup>5</sup>SGKL<sup>6</sup>DE<sup>7</sup>WE<sup>8</sup>KIRL<sup>9</sup>RP<sup>10</sup>GG<sup>11</sup>KK<sup>12</sup>YRL<sup>13</sup>KL<sup>14</sup>VW<sup>15</sup>AS<sup>16</sup>RE<sup>17</sup>LE<sup>18</sup>RF<sup>19</sup>AL<sup>20</sup>NP<sup>21</sup>GL<sup>22</sup>LE<sup>23</sup>TA<sup>24</sup>EG<sup>25</sup>CQ<sup>26</sup>Q<sup>27</sup>IE<sup>28</sup>QL<sup>29</sup>SA<sup>30</sup>LK<sup>31</sup>TG<sup>32</sup>SE<sup>33</sup>EL<sup>34</sup>KS<sup>35</sup>LY<sup>36</sup>NT<sup>37</sup>VAT<sup>38</sup>LY<sup>39</sup>CV<sup>40</sup>HQ<sup>41</sup>RI<sup>42</sup>  
 K<sup>43</sup>VT<sup>44</sup>DT<sup>45</sup>KE<sup>46</sup>AL<sup>47</sup>DK<sup>48</sup>IE<sup>49</sup>EQ<sup>50</sup>NS<sup>51</sup>Q<sup>52</sup>Y<sup>53</sup>PI<sup>54</sup>VQ<sup>55</sup>NA<sup>56</sup>Q<sup>57</sup>Q<sup>58</sup>VM<sup>59</sup>HQ<sup>60</sup>AI<sup>61</sup>SP<sup>62</sup>RL<sup>63</sup>NA<sup>64</sup>W<sup>65</sup>K<sup>66</sup>VI<sup>67</sup>E<sup>68</sup>E<sup>69</sup>KA<sup>70</sup>FS<sup>71</sup>PE<sup>72</sup>VI<sup>73</sup>PM<sup>74</sup>FS<sup>75</sup>AL<sup>76</sup>SE<sup>77</sup>GAT<sup>78</sup>PP<sup>79</sup>Q<sup>80</sup>DL<sup>81</sup>  
 NM<sup>82</sup>LN<sup>83</sup>IV<sup>84</sup>GG<sup>85</sup>HQ<sup>86</sup>AA<sup>87</sup>QM<sup>88</sup>KL<sup>89</sup>DT<sup>90</sup>IN<sup>91</sup>EE<sup>92</sup>AE<sup>93</sup>W<sup>94</sup>DR<sup>95</sup>VP<sup>96</sup>HA<sup>97</sup>GI<sup>98</sup>PP<sup>99</sup>Q<sup>100</sup>RE<sup>101</sup>PR<sup>102</sup>GS<sup>103</sup>DI<sup>104</sup>AG<sup>105</sup>TT<sup>106</sup>SL<sup>107</sup>Q<sup>108</sup>EQ<sup>109</sup>IG<sup>110</sup>W<sup>111</sup>TS<sup>112</sup>NP<sup>113</sup>PI<sup>114</sup>PV<sup>115</sup>GE<sup>116</sup>IY<sup>117</sup>K<sup>118</sup>R<sup>119</sup>WI<sup>120</sup>IL<sup>121</sup>GL<sup>122</sup>NK<sup>123</sup>IV<sup>124</sup>RM<sup>125</sup>  
 Y<sup>126</sup>SP<sup>127</sup>VS<sup>128</sup>IL<sup>129</sup>DI<sup>130</sup>RQ<sup>131</sup>G<sup>132</sup>PK<sup>133</sup>EP<sup>134</sup>FR<sup>135</sup>DY<sup>136</sup>VD<sup>137</sup>RF<sup>138</sup>FK<sup>139</sup>TL<sup>140</sup>RA<sup>141</sup>EQ<sup>142</sup>AT<sup>143</sup>Q<sup>144</sup>EV<sup>145</sup>KN<sup>146</sup>WT<sup>147</sup>DT<sup>148</sup>LL<sup>149</sup>VQ<sup>150</sup>NA<sup>151</sup>NP<sup>152</sup>DC<sup>153</sup>KT<sup>154</sup>IL<sup>155</sup>KAL<sup>156</sup>GP<sup>157</sup>AT<sup>158</sup>LE<sup>159</sup>EM<sup>160</sup>MT<sup>161</sup>AC<sup>162</sup>Q<sup>163</sup>GV<sup>164</sup>GG<sup>165</sup>PG<sup>166</sup>HK<sup>167</sup>AR<sup>168</sup>VL<sup>169</sup>AE<sup>170</sup>AMS<sup>171</sup>  
 QAS<sup>172</sup>GT<sup>173</sup>EA<sup>174</sup>AIM<sup>175</sup>MQ<sup>176</sup>KS<sup>177</sup>N<sup>178</sup>FK<sup>179</sup>GP<sup>180</sup>KRS<sup>181</sup>IK<sup>182</sup>CF<sup>183</sup>NC<sup>184</sup>GK<sup>185</sup>EG<sup>186</sup>HL<sup>187</sup>ARN<sup>188</sup>CR<sup>189</sup>AP<sup>190</sup>RR<sup>191</sup>KK<sup>192</sup>GC<sup>193</sup>WK<sup>194</sup>CK<sup>195</sup>EG<sup>196</sup>HQ<sup>197</sup>MK<sup>198</sup>DC<sup>199</sup>TER<sup>200</sup>QAN<sup>201</sup>FL<sup>202</sup>KI<sup>203</sup>W<sup>204</sup>PS<sup>205</sup>NK<sup>206</sup>GR<sup>207</sup>PN<sup>208</sup>GL<sup>209</sup>Q<sup>210</sup>NR<sup>211</sup>PE<sup>212</sup>PT<sup>213</sup>APP<sup>214</sup>  
 AES<sup>215</sup>FG<sup>216</sup>GE<sup>217</sup>ET<sup>218</sup>AP<sup>219</sup>SP<sup>220</sup>KQ<sup>221</sup>EP<sup>222</sup>KE<sup>223</sup>KE<sup>224</sup>LY<sup>225</sup>PL<sup>226</sup>AS<sup>227</sup>LK<sup>228</sup>SL<sup>229</sup>FG<sup>230</sup>ND<sup>231</sup>P<sup>232</sup>‡

Fig. 78B

## 2003\_CON\_06\_CPX\_gag.OPT

AT<sup>1</sup>GG<sup>2</sup>CG<sup>3</sup>C<sup>4</sup>CG<sup>5</sup>C<sup>6</sup>CG<sup>7</sup>C<sup>8</sup>CG<sup>9</sup>C<sup>10</sup>CG<sup>11</sup>C<sup>12</sup>CG<sup>13</sup>C<sup>14</sup>CG<sup>15</sup>C<sup>16</sup>CG<sup>17</sup>C<sup>18</sup>CG<sup>19</sup>C<sup>20</sup>CG<sup>21</sup>C<sup>22</sup>CG<sup>23</sup>C<sup>24</sup>CG<sup>25</sup>C<sup>26</sup>CG<sup>27</sup>C<sup>28</sup>CG<sup>29</sup>C<sup>30</sup>CG<sup>31</sup>C<sup>32</sup>CG<sup>33</sup>C<sup>34</sup>CG<sup>35</sup>C<sup>36</sup>CG<sup>37</sup>C<sup>38</sup>CG<sup>39</sup>C<sup>40</sup>CG<sup>41</sup>C<sup>42</sup>CG<sup>43</sup>C<sup>44</sup>CG<sup>45</sup>C<sup>46</sup>CG<sup>47</sup>C<sup>48</sup>CG<sup>49</sup>C<sup>50</sup>CG<sup>51</sup>C<sup>52</sup>CG<sup>53</sup>C<sup>54</sup>CG<sup>55</sup>C<sup>56</sup>CG<sup>57</sup>C<sup>58</sup>CG<sup>59</sup>C<sup>60</sup>CG<sup>61</sup>C<sup>62</sup>CG<sup>63</sup>C<sup>64</sup>CG<sup>65</sup>C<sup>66</sup>CG<sup>67</sup>C<sup>68</sup>CG<sup>69</sup>C<sup>70</sup>CG<sup>71</sup>C<sup>72</sup>CG<sup>73</sup>C<sup>74</sup>CG<sup>75</sup>C<sup>76</sup>CG<sup>77</sup>C<sup>78</sup>CG<sup>79</sup>C<sup>80</sup>CG<sup>81</sup>C<sup>82</sup>CG<sup>83</sup>C<sup>84</sup>CG<sup>85</sup>C<sup>86</sup>CG<sup>87</sup>C<sup>88</sup>CG<sup>89</sup>C<sup>90</sup>CG<sup>91</sup>C<sup>92</sup>CG<sup>93</sup>C<sup>94</sup>CG<sup>95</sup>C<sup>96</sup>CG<sup>97</sup>C<sup>98</sup>CG<sup>99</sup>C<sup>100</sup>CG<sup>101</sup>C<sup>102</sup>CG<sup>103</sup>C<sup>104</sup>CG<sup>105</sup>C<sup>106</sup>CG<sup>107</sup>C<sup>108</sup>CG<sup>109</sup>C<sup>110</sup>CG<sup>111</sup>C<sup>112</sup>CG<sup>113</sup>C<sup>114</sup>CG<sup>115</sup>C<sup>116</sup>CG<sup>117</sup>C<sup>118</sup>CG<sup>119</sup>C<sup>120</sup>CG<sup>121</sup>C<sup>122</sup>CG<sup>123</sup>C<sup>124</sup>CG<sup>125</sup>C<sup>126</sup>CG<sup>127</sup>C<sup>128</sup>CG<sup>129</sup>C<sup>130</sup>CG<sup>131</sup>C<sup>132</sup>CG<sup>133</sup>C<sup>134</sup>CG<sup>135</sup>C<sup>136</sup>CG<sup>137</sup>C<sup>138</sup>CG<sup>139</sup>C<sup>140</sup>CG<sup>141</sup>C<sup>142</sup>CG<sup>143</sup>C<sup>144</sup>CG<sup>145</sup>C<sup>146</sup>CG<sup>147</sup>C<sup>148</sup>CG<sup>149</sup>C<sup>150</sup>CG<sup>151</sup>C<sup>152</sup>CG<sup>153</sup>C<sup>154</sup>CG<sup>155</sup>C<sup>156</sup>CG<sup>157</sup>C<sup>158</sup>CG<sup>159</sup>C<sup>160</sup>CG<sup>161</sup>C<sup>162</sup>CG<sup>163</sup>C<sup>164</sup>CG<sup>165</sup>C<sup>166</sup>CG<sup>167</sup>C<sup>168</sup>CG<sup>169</sup>C<sup>170</sup>CG<sup>171</sup>C<sup>172</sup>CG<sup>173</sup>C<sup>174</sup>CG<sup>175</sup>C<sup>176</sup>CG<sup>177</sup>C<sup>178</sup>CG<sup>179</sup>C<sup>180</sup>CG<sup>181</sup>C<sup>182</sup>CG<sup>183</sup>C<sup>184</sup>CG<sup>185</sup>C<sup>186</sup>CG<sup>187</sup>C<sup>188</sup>CG<sup>189</sup>C<sup>190</sup>CG<sup>191</sup>C<sup>192</sup>CG<sup>193</sup>C<sup>194</sup>CG<sup>195</sup>C<sup>196</sup>CG<sup>197</sup>C<sup>198</sup>CG<sup>199</sup>C<sup>200</sup>CG<sup>201</sup>C<sup>202</sup>CG<sup>203</sup>C<sup>204</sup>CG<sup>205</sup>C<sup>206</sup>CG<sup>207</sup>C<sup>208</sup>CG<sup>209</sup>C<sup>210</sup>CG<sup>211</sup>C<sup>212</sup>CG<sup>213</sup>C<sup>214</sup>CG<sup>215</sup>C<sup>216</sup>CG<sup>217</sup>C<sup>218</sup>CG<sup>219</sup>C<sup>220</sup>CG<sup>221</sup>C<sup>222</sup>CG<sup>223</sup>C<sup>224</sup>CG<sup>225</sup>C<sup>226</sup>CG<sup>227</sup>C<sup>228</sup>CG<sup>229</sup>C<sup>230</sup>CG<sup>231</sup>C<sup>232</sup>CG<sup>233</sup>C<sup>234</sup>CG<sup>235</sup>C<sup>236</sup>CG<sup>237</sup>C<sup>238</sup>CG<sup>239</sup>C<sup>240</sup>CG<sup>241</sup>C<sup>242</sup>CG<sup>243</sup>C<sup>244</sup>CG<sup>245</sup>C<sup>246</sup>CG<sup>247</sup>C<sup>248</sup>CG<sup>249</sup>C<sup>250</sup>CG<sup>251</sup>C<sup>252</sup>CG<sup>253</sup>C<sup>254</sup>CG<sup>255</sup>C<sup>256</sup>CG<sup>257</sup>C<sup>258</sup>CG<sup>259</sup>C<sup>260</sup>CG<sup>261</sup>C<sup>262</sup>CG<sup>263</sup>C<sup>264</sup>CG<sup>265</sup>C<sup>266</sup>CG<sup>267</sup>C<sup>268</sup>CG<sup>269</sup>C<sup>270</sup>CG<sup>271</sup>C<sup>272</sup>CG<sup>273</sup>C<sup>274</sup>CG<sup>275</sup>C<sup>276</sup>CG<sup>277</sup>C<sup>278</sup>CG<sup>279</sup>C<sup>280</sup>CG<sup>281</sup>C<sup>282</sup>CG<sup>283</sup>C<sup>284</sup>CG<sup>285</sup>C<sup>286</sup>CG<sup>287</sup>C<sup>288</sup>CG<sup>289</sup>C<sup>290</sup>CG<sup>291</sup>C<sup>292</sup>CG<sup>293</sup>C<sup>294</sup>CG<sup>295</sup>C<sup>296</sup>CG<sup>297</sup>C<sup>298</sup>CG<sup>299</sup>C<sup>300</sup>CG<sup>301</sup>C<sup>302</sup>CG<sup>303</sup>C<sup>304</sup>CG<sup>305</sup>C<sup>306</sup>CG<sup>307</sup>C<sup>308</sup>CG<sup>309</sup>C<sup>310</sup>CG<sup>311</sup>C<sup>312</sup>CG<sup>313</sup>C<sup>314</sup>CG<sup>315</sup>C<sup>316</sup>CG<sup>317</sup>C<sup>318</sup>CG<sup>319</sup>C<sup>320</sup>CG<sup>321</sup>C<sup>322</sup>CG<sup>323</sup>C<sup>324</sup>CG<sup>325</sup>C<sup>326</sup>CG<sup>327</sup>C<sup>328</sup>CG<sup>329</sup>C<sup>330</sup>CG<sup>331</sup>C<sup>332</sup>CG<sup>333</sup>C<sup>334</sup>CG<sup>335</sup>C<sup>336</sup>CG<sup>337</sup>C<sup>338</sup>CG<sup>339</sup>C<sup>340</sup>CG<sup>341</sup>C<sup>342</sup>CG<sup>343</sup>C<sup>344</sup>CG<sup>345</sup>C<sup>346</sup>CG<sup>347</sup>C<sup>348</sup>CG<sup>349</sup>C<sup>350</sup>CG<sup>351</sup>C<sup>352</sup>CG<sup>353</sup>C<sup>354</sup>CG<sup>355</sup>C<sup>356</sup>CG<sup>357</sup>C<sup>358</sup>CG<sup>359</sup>C<sup>360</sup>CG<sup>361</sup>C<sup>362</sup>CG<sup>363</sup>C<sup>364</sup>CG<sup>365</sup>C<sup>366</sup>CG<sup>367</sup>C<sup>368</sup>CG<sup>369</sup>C<sup>370</sup>CG<sup>371</sup>C<sup>372</sup>CG<sup>373</sup>C<sup>374</sup>CG<sup>375</sup>C<sup>376</sup>CG<sup>377</sup>C<sup>378</sup>CG<sup>379</sup>C<sup>380</sup>CG<sup>381</sup>C<sup>382</sup>CG<sup>383</sup>C<sup>384</sup>CG<sup>385</sup>C<sup>386</sup>CG<sup>387</sup>C<sup>388</sup>CG<sup>389</sup>C<sup>390</sup>CG<sup>391</sup>C<sup>392</sup>CG<sup>393</sup>C<sup>394</sup>CG<sup>395</sup>C<sup>396</sup>CG<sup>397</sup>C<sup>398</sup>CG<sup>399</sup>C<sup>400</sup>CG<sup>401</sup>C<sup>402</sup>CG<sup>403</sup>C<sup>404</sup>CG<sup>405</sup>C<sup>406</sup>CG<sup>407</sup>C<sup>408</sup>CG<sup>409</sup>C<sup>410</sup>CG<sup>411</sup>C<sup>412</sup>CG<sup>413</sup>C<sup>414</sup>CG<sup>415</sup>C<sup>416</sup>CG<sup>417</sup>C<sup>418</sup>CG<sup>419</sup>C<sup>420</sup>CG<sup>421</sup>C<sup>422</sup>CG<sup>423</sup>C<sup>424</sup>CG<sup>425</sup>C<sup>426</sup>CG<sup>427</sup>C<sup>428</sup>CG<sup>429</sup>C<sup>430</sup>CG<sup>431</sup>C<sup>432</sup>CG<sup>433</sup>C<sup>434</sup>CG<sup>435</sup>C<sup>436</sup>CG<sup>437</sup>C<sup>438</sup>CG<sup>439</sup>C<sup>440</sup>CG<sup>441</sup>C<sup>442</sup>CG<sup>443</sup>C<sup>444</sup>CG<sup>445</sup>C<sup>446</sup>CG<sup>447</sup>C<sup>448</sup>CG<sup>449</sup>C<sup>450</sup>CG<sup>451</sup>C<sup>452</sup>CG<sup>453</sup>C<sup>454</sup>CG<sup>455</sup>C<sup>456</sup>CG<sup>457</sup>C<sup>458</sup>CG<sup>459</sup>C<sup>460</sup>CG<sup>461</sup>C<sup>462</sup>CG<sup>463</sup>C<sup>464</sup>CG<sup>465</sup>C<sup>466</sup>CG<sup>467</sup>C<sup>468</sup>CG<sup>469</sup>C<sup>470</sup>CG<sup>471</sup>C<sup>472</sup>CG<sup>473</sup>C<sup>474</sup>CG<sup>475</sup>C<sup>476</sup>CG<sup>477</sup>C<sup>478</sup>CG<sup>479</sup>C<sup>480</sup>CG<sup>481</sup>C<sup>482</sup>CG<sup>483</sup>C<sup>484</sup>CG<sup>485</sup>C<sup>486</sup>CG<sup>487</sup>C<sup>488</sup>CG<sup>489</sup>C<sup>490</sup>CG<sup>491</sup>C<sup>492</sup>CG<sup>493</sup>C<sup>494</sup>CG<sup>495</sup>C<sup>496</sup>CG<sup>497</sup>C<sup>498</sup>CG<sup>499</sup>C<sup>500</sup>CG<sup>501</sup>C<sup>502</sup>CG<sup>503</sup>C<sup>504</sup>CG<sup>505</sup>C<sup>506</sup>CG<sup>507</sup>C<sup>508</sup>CG<sup>509</sup>C<sup>510</sup>CG<sup>511</sup>C<sup>512</sup>CG<sup>513</sup>C<sup>514</sup>CG<sup>515</sup>C<sup>516</sup>CG<sup>517</sup>C<sup>518</sup>CG<sup>519</sup>C<sup>520</sup>CG<sup>521</sup>C<sup>522</sup>CG<sup>523</sup>C<sup>524</sup>CG<sup>525</sup>C<sup>526</sup>CG<sup>527</sup>C<sup>528</sup>CG<sup>529</sup>C<sup>530</sup>CG<sup>531</sup>C<sup>532</sup>CG<sup>533</sup>C<sup>534</sup>CG<sup>535</sup>C<sup>536</sup>CG<sup>537</sup>C<sup>538</sup>CG<sup>539</sup>C<sup>540</sup>CG<sup>541</sup>C<sup>542</sup>CG<sup>543</sup>C<sup>544</sup>CG<sup>545</sup>C<sup>546</sup>CG<sup>547</sup>C<sup>548</sup>CG<sup>549</sup>C<sup>550</sup>CG<sup>551</sup>C<sup>552</sup>CG<sup>553</sup>C<sup>554</sup>CG<sup>555</sup>C<sup>556</sup>CG<sup>557</sup>C<sup>558</sup>CG<sup>559</sup>C<sup>560</sup>CG<sup>561</sup>C<sup>562</sup>CG<sup>563</sup>C<sup>564</sup>CG<sup>565</sup>C<sup>566</sup>CG<sup>567</sup>C<sup>568</sup>CG<sup>569</sup>C<sup>570</sup>CG<sup>571</sup>C<sup>572</sup>CG<sup>573</sup>C<sup>574</sup>CG<sup>575</sup>C<sup>576</sup>CG<sup>577</sup>C<sup>578</sup>CG<sup>579</sup>C<sup>580</sup>CG<sup>581</sup>C<sup>582</sup>CG<sup>583</sup>C<sup>584</sup>CG<sup>585</sup>C<sup>586</sup>CG<sup>587</sup>C<sup>588</sup>CG<sup>589</sup>C<sup>590</sup>CG<sup>591</sup>C<sup>592</sup>CG<sup>593</sup>C<sup>594</sup>CG<sup>595</sup>C<sup>596</sup>CG<sup>597</sup>C<sup>598</sup>CG<sup>599</sup>C<sup>600</sup>CG<sup>601</sup>C<sup>602</sup>CG<sup>603</sup>C<sup>604</sup>CG<sup>605</sup>C<sup>606</sup>CG<sup>607</sup>C<sup>608</sup>CG<sup>609</sup>C<sup>610</sup>CG<sup>611</sup>C<sup>612</sup>CG<sup>613</sup>C<sup>614</sup>CG<sup>615</sup>C<sup>616</sup>CG<sup>617</sup>C<sup>618</sup>CG<sup>619</sup>C<sup>620</sup>CG<sup>621</sup>C<sup>622</sup>CG<sup>623</sup>C<sup>624</sup>CG<sup>625</sup>C<sup>626</sup>CG<sup>627</sup>C<sup>628</sup>CG<sup>629</sup>C<sup>630</sup>CG<sup>631</sup>C<sup>632</sup>CG<sup>633</sup>C<sup>634</sup>CG<sup>635</sup>C<sup>636</sup>CG<sup>637</sup>C<sup>638</sup>CG<sup>639</sup>C<sup>640</sup>CG<sup>641</sup>C<sup>642</sup>CG<sup>643</sup>C<sup>644</sup>CG<sup>645</sup>C<sup>646</sup>CG<sup>647</sup>C<sup>648</sup>CG<sup>649</sup>C<sup>650</sup>CG<sup>651</sup>C<sup>652</sup>CG<sup>653</sup>C<sup>654</sup>CG<sup>655</sup>C<sup>656</sup>CG<sup>657</sup>C<sup>658</sup>CG<sup>659</sup>C<sup>660</sup>CG<sup>661</sup>C<sup>662</sup>CG<sup>663</sup>C<sup>664</sup>CG<sup>665</sup>C<sup>666</sup>CG<sup>667</sup>C<sup>668</sup>CG<sup>669</sup>C<sup>670</sup>CG<sup>671</sup>C<sup>672</sup>CG<sup>673</sup>C<sup>674</sup>CG<sup>675</sup>C<sup>676</sup>CG<sup>677</sup>C<sup>678</sup>CG<sup>679</sup>C<sup>680</sup>CG<sup>681</sup>C<sup>682</sup>CG<sup>683</sup>C<sup>684</sup>CG<sup>685</sup>C<sup>686</sup>CG<sup>687</sup>C<sup>688</sup>CG<sup>689</sup>C<sup>690</sup>CG<sup>691</sup>C<sup>692</sup>CG<sup>693</sup>C<sup>694</sup>CG<sup>695</sup>C<sup>696</sup>CG<sup>697</sup>C<sup>698</sup>CG<sup>699</sup>C<sup>700</sup>CG<sup>701</sup>C<sup>702</sup>CG<sup>703</sup>C<sup>704</sup>CG<sup>705</sup>C<sup>706</sup>CG<sup>707</sup>C<sup>708</sup>CG<sup>709</sup>C<sup>710</sup>CG<sup>711</sup>C<sup>712</sup>CG<sup>713</sup>C<sup>714</sup>CG<sup>715</sup>C<sup>716</sup>CG<sup>717</sup>C<sup>718</sup>CG<sup>719</sup>C<sup>720</sup>CG<sup>721</sup>C<sup>722</sup>CG<sup>723</sup>C<sup>724</sup>CG<sup>725</sup>C<sup>726</sup>CG<sup>727</sup>C<sup>728</sup>CG<sup>729</sup>C<sup>730</sup>CG<sup>731</sup>C<sup>732</sup>CG<sup>733</sup>C<sup>734</sup>CG<sup>735</sup>C<sup>736</sup>CG<sup>737</sup>C<sup>738</sup>CG<sup>739</sup>C<sup>740</sup>CG<sup>741</sup>C<sup>742</sup>CG<sup>743</sup>C<sup>744</sup>CG<sup>745</sup>C<sup>746</sup>CG<sup>747</sup>C<sup>748</sup>CG<sup>749</sup>C<sup>750</sup>CG<sup>751</sup>C<sup>752</sup>CG<sup>753</sup>C<sup>754</sup>CG<sup>755</sup>C<sup>756</sup>CG<sup>757</sup>C<sup>758</sup>CG<sup>759</sup>C<sup>760</sup>CG<sup>761</sup>C<sup>762</sup>CG<sup>763</sup>C<sup>764</sup>CG<sup>765</sup>C<sup>766</sup>CG<sup>767</sup>C<sup>768</sup>CG<sup>769</sup>C<sup>770</sup>CG<sup>771</sup>C<sup>772</sup>CG<sup>773</sup>C<sup>774</sup>CG<sup>775</sup>C<sup>776</sup>CG<sup>777</sup>C<sup>778</sup>CG<sup>779</sup>C<sup>780</sup>CG<sup>781</sup>C<sup>782</sup>CG<sup>783</sup>C<sup>784</sup>CG<sup>785</sup>C<sup>786</sup>CG<sup>787</sup>C<sup>788</sup>CG<sup>789</sup>C<sup>790</sup>CG<sup>791</sup>C<sup>792</sup>CG<sup>793</sup>C<sup>794</sup>CG<sup>795</sup>C<sup>796</sup>CG<sup>797</sup>C<sup>798</sup>CG<sup>799</sup>C<sup>800</sup>CG<sup>801</sup>C<sup>802</sup>CG<sup>803</sup>C<sup>804</sup>CG<sup>805</sup>C<sup>806</sup>CG<sup>807</sup>C<sup>808</sup>CG<sup>809</sup>C<sup>810</sup>CG<sup>811</sup>C<sup>812</sup>CG<sup>813</sup>C<sup>814</sup>CG<sup>815</sup>C<sup>816</sup>CG<sup>817</sup>C<sup>818</sup>CG<sup>819</sup>C<sup>820</sup>CG<sup>821</sup>C<sup>822</sup>CG<sup>823</sup>C<sup>824</sup>CG<sup>825</sup>C<sup>826</sup>CG<sup>827</sup>C<sup>828</sup>CG<sup>829</sup>C<sup>830</sup>CG<sup>831</sup>C<sup>832</sup>CG<sup>833</sup>C<sup>834</sup>CG<sup>835</sup>C<sup>836</sup>CG<sup>837</sup>C<sup>838</sup>CG<sup>839</sup>C<sup>840</sup>CG<sup>841</sup>C<sup>842</sup>CG<sup>843</sup>C<sup>844</sup>CG<sup>845</sup>C<sup>846</sup>CG<sup>847</sup>C<sup>848</sup>CG<sup>849</sup>C<sup>850</sup>CG<sup>851</sup>C<sup>852</sup>CG<sup>853</sup>C<sup>854</sup>CG<sup>855</sup>C<sup>856</sup>CG<sup>857</sup>C<sup>858</sup>CG<sup>859</sup>C<sup>860</sup>CG<sup>861</sup>C<sup>862</sup>CG<sup>863</sup>C<sup>864</sup>CG<sup>865</sup>C<sup>866</sup>CG<sup>867</sup>C<sup>868</sup>CG<sup>869</sup>C<sup>870</sup>CG<sup>871</sup>C<sup>872</sup>CG<sup>873</sup>C<sup>874</sup>CG<sup>875</sup>C<sup>876</sup>CG<sup>877</sup>C<sup>878</sup>CG<sup>879</sup>C<sup>880</sup>CG<sup>881</sup>C<sup>882</sup>CG<sup>883</sup>C<sup>884</sup>CG<sup>885</sup>C<sup>886</sup>CG<sup>887</sup>C<sup>888</sup>CG<sup>889</sup>C<sup>890</sup>CG<sup>891</sup>C<sup>892</sup>CG<sup>893</sup>C<sup>894</sup>CG<sup>895</sup>C<sup>896</sup>CG<sup>897</sup>C<sup>898</sup>CG<sup>899</sup>C<sup>900</sup>CG<sup>901</sup>C<sup>902</sup>CG<sup>903</sup>C<sup>904</sup>CG<sup>905</sup>C<sup>906</sup>CG<sup>907</sup>C<sup>908</sup>CG<sup>909</sup>C<sup>910</sup>CG<sup>911</sup>C<sup>912</sup>CG<sup>913</sup>C<sup>914</sup>CG<sup>915</sup>C<sup>916</sup>CG<sup>917</sup>C<sup>918</sup>CG<sup>919</sup>C<sup>920</sup>CG<sup>921</sup>C<sup>922</sup>CG<sup>923</sup>C<sup>924</sup>CG<sup>925</sup>C<sup>926</sup>CG<sup>927</sup>C<sup>928</sup>CG<sup>929</sup>C<sup>930</sup>CG<sup>931</sup>C<sup>932</sup>CG<sup>933</sup>C<sup>934</sup>CG<sup>935</sup>C<sup>936</sup>CG<sup>937</sup>C<sup>938</sup>CG<sup>939</sup>C<sup>940</sup>CG<sup>941</sup>C<sup>942</sup>CG<sup>943</sup>C<sup>944</sup>CG<sup>945</sup>C<sup>946</sup>CG<sup>947</sup>C<sup>948</sup>CG<sup>949</sup>C<sup>950</sup>CG<sup>951</sup>C<sup>952</sup>CG<sup>953</sup>C<sup>954</sup>CG<sup>955</sup>C<sup>956</sup>CG<sup>957</sup>C<sup>958</sup>CG<sup>959</sup>C<sup>960</sup>CG<sup>961</sup>C<sup>962</sup>CG<sup>963</sup>C<sup>964</sup>CG<sup>965</sup>C<sup>966</sup>CG<sup>967</sup>C<sup>968</sup>CG<sup>969</sup>C<sup>970</sup>CG<sup>971</sup>C<sup>972</sup>CG<sup>973</sup>C<sup>974</sup>CG<sup>975</sup>C<sup>976</sup>CG<sup>977</sup>C<sup>978</sup>CG<sup>979</sup>C<sup>980</sup>CG<sup>981</sup>C<sup>982</sup>CG<sup>983</sup>C<sup>984</sup>CG<sup>985</sup>C<sup>986</sup>CG<sup>987</sup>C<sup>988</sup>CG<sup>989</sup>C<sup>990</sup>CG<sup>991</sup>C<sup>992</sup>CG<sup>993</sup>C<sup>994</sup>CG<sup>995</sup>C<sup>996</sup>CG<sup>997</sup>C<sup>998</sup>CG<sup>999</sup>C<sup>1000</sup>CG<sup>1001</sup>C<sup>1002</sup>CG<sup>1003</sup>C<sup>1004</sup>CG<sup>1005</sup>C<sup>1006</sup>CG<sup>1007</sup>C<sup>1008</sup>CG<sup>1009</sup>C<sup>1010</sup>CG<sup>1011</sup>C<sup>1012</sup>CG<sup>1013</sup>C<sup>1014</sup>CG<sup>1015</sup>C<sup>1016</sup>CG<sup>1017</sup>C<sup>1018</sup>CG<sup>1019</sup>C<sup>1020</sup>CG<sup>1021</sup>C<sup>1022</sup>CG<sup>1023</sup>C<sup>1024</sup>CG<sup>1025</sup>C<sup>1026</sup>CG<sup>1027</sup>C<sup>1028</sup>CG<sup>1029</sup>C<sup>1030</sup>CG<sup>1031</sup>C<sup>1032</sup>CG<sup>1033</sup>C<sup>1034</sup>CG<sup>1035</sup>C<sup>1036</sup>CG<sup>1037</sup>C<sup>1038</sup>CG<sup>1039</sup>C<sup>1040</sup>CG<sup>1041</sup>C<sup>1042</sup>CG<sup>1043</sup>C<sup>1044</sup>CG<

122/178

Fig. 79A

20. 2003 CON 07 BC gag. PEP  
 MGARASILRGGLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIIKQLPALQTGTEELRSLFNTVATLYCVHTEI  
 DVRTKEALDKIEEEQNKIQQKTQQAKEADGKVSQNYPIVQNLQGMVHQPISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTM  
 INTVGGHQAAQIILKDTINEEAAEWDRLHPVHAGPIAPGQMRPRGSDIAGTTSNLQEQIAWMTSNPPVPVGDIIYKRWIIILGNKIVRMYS  
 TSILDIKQGPKEFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNPANPDCKTILRALPGASIEEMMTACQGVGGPSHKARVIAEAMSQTN  
 STILMORSNFKGSKRIVKCFNCGKEGHIARNCRAPRKKGCKGKEGHQMKDCTERQANFLGKIWP SHKGRPNFLQSRPEPTAPPEESFRF  
 GEETTPSQKQEPIDKELYPLTSLKSLFGNDPSSQ\$

Fig. 79B

2003 CON 07 BC gag. OPT  
 ATGGGCGCCCGGCTCCATCCTGCGGGGGCAAGCTGGACAAGTGGGAGAAGATCCGCCCTGCGCCCCGGGGCAAGAAAGCACTACATGCT  
 GAAGCACCTGGTGTGGCCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCTGTCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA  
 AGCAGCTGCAGCCCGCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTTCAACACCGTGGCCACCTGTACTGCGTGCACACCGAGATC  
 GACGTGCGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGACAAGATCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGACGG  
 CAAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGACACAGCCCATCTCCCCCGCACCCCTGAACGCCCTGGGTGA  
 AGTGGTGGAGGAGAAGCCCTTCCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGAGTGGGACCCGCTGCACCCCGT  
 CTGAACACCGTGGGCGCCACCAAGCGCCATGCAGATCCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGGACCCGCTGCACCCCGT  
 GCACGCGGCCCATCGCCCCGCGCAGATGCGCGAGCCCCCGGCTCCGACATCGCGGCACCACTCCAACCTGCAGGAGCAGATCGCCT  
 GGATGACCTCCAACCCCGTGGCGGACATCTACAAGCGCTGGATCATCCTGGGCCCTGAACAAGATCGTGGCATGTACTCCCCC  
 ACCTCCATCTTGACATCAAGCAGGGCCCCAAGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCTCGCGCGCCGAGCAGGCCAC  
 CCAGGACGTGAAGAACTGGATGACCGACACCTTGCTGGTGCAGAACGCCAACCCTGCTGAGACCTCAAGACCATCTGCGCGCCCTGGGCCCGGCG  
 CCTCCATCGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCACAAGGCCCGCTGTGGCCGAGGCCATGTCCCAGAACCAAC  
 TCCACCATCTGATGCAGCGCTCCAACTTCAAGGCTCCAAGCGCATCGTGAAGTCTTCAACTGGGGCAAGGAGGCCACATCGCCCCGCAA  
 CTGCCGCGCCCCCGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTGTTGG  
 GCAAGATCTGGCCCTCCACAAAGGGCGGCCCGCAACTTCTGAGTCCCGCCCGAGCCACCGCCCGGAGGAGTCTTCCGCTTC  
 GCGAGGAGACCAACCCCTCCAGAAAGCAGGAGGCCCATCGACAAGGAGCTGTACCCCTGACCTCCCTGAAAGTCCCTGTTCGGCAACGA  
 CCCCCTCCCCAGTAA

123/178

Fig. 80A

## 21. 2003 CON 08 BC gag. PEP

MGARASILRGGLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIIKQLPALQGTTEELRSLFNTVATLYCVHAEI  
 EVRDTKEALDKIEEEQNKIQKTQQAKEADEKVSQNYPIVQNLQGQMVHQPLSPRTLNAWVKVVEEKAFSPEVIPMFALTALSEGATPQDLNTM  
 LNTVGGHQAAQMMLKDTINEEAAEWDRLHPVHAGPVAPGQMPREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYS  
 TSILDIKQGPKEPRFDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNPANPDCKTILRALGPGASLEEMMTACQGVGPPSHKARVLAEAMSQTN  
 NTILMORSNFKSKRIVKCFNCGKEGHIANKRPRKKGCKGKEGHQMKDCTERQANFLGKIWP SHKGRPGNFIQSRPEPTAPPAESFRF  
 EETTPAPKQEPKDRPLETSLRSLFGSDPLSQ\$

Fig. 80B

## 2003 CON 08 BC gag. OPT

ATGGCGCGCGCGCTCCATCTGCGCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGGCTCGGCCCGGCGGCAAGAAGCACTACATGCT  
 GAAGCACTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCGGCTGTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA  
 AGCAGCTGCAGCCCGCTGCAGACCGGCACCGAGGAGCTGGCTCCCTGTTCAACACCGTGGCCACCTGTACTGCGTGCAAGCCGAGATC  
 GAGGTGCGGCACACCAAGGAGCCCTGGACAAGATCGAGGAGGAGAGCAACAAGATCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGACGA  
 GAAGGTGTCCAGAACTACCCCATCTGTGCAAACTGTGACAGCCCTGTCCAGCCCTGTCCAGGGCGCACCCAGGACCTGAACGCTGGGTGA  
 AGGTGGTGGAGGAGAGGCTTCTCCCCGAGGTGATCCCCATGTTACCGCCCTGTCCAGGGCGCACCCAGGACCTGAACACCATG  
 CTGAACACCGTGGCGGCCACAGGCCCGCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCGAGTGGACCCCTGCACCCCGT  
 GCACGCCGCCCGTGGCCCCCGCCAGATGCGCGAGCCCGCGGCTCCGACATCGCCGGCACCATCCACCTGCAGGAGCAGATCGGCT  
 GGATGACCAACAAACCCCATCCCCGTGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCCTGAACAAGATCGTGCGCATGTACTCCCC  
 ACCTCCATCCTGGACATCAAGCAGGCCCAAGAGCCCTCCCGGACTACGTGACCGCTTCTTCAAGACCTTCGCGCCGAGCAGGCCAC  
 CCAGGACGTGAAGAACTGGATGACCGACACCCCTGCTGGTGCAGAACGCCAACCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCGGCG  
 CCTCCCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCACAAGGCCCGCTGCTGGCCGAGGCCATGTCCAGACCAAC  
 AACACCATCCTGATGCAGCGCTCCAACTTCAAGGGCTCCAAGCGCATCGTGAAGTCTTCAACTGCGGCAAGGAGGCCACATCGCCAGAA  
 CTGCGCGGCCCGGCAAGAGGCTGCTGGAAGTGGGCAAGGAGGCCACCATGATGAAGGACTGCACCGAGGCCAGGCCAACTTCTGTGG  
 GCAAGATCTGGCCCTCCCAAGAGGCCCGCCCGCAACTTCTGTGAGTCCCGCCCGAGCCCAACCGCCCCCGGAGTCTTCCGCTTC  
 GAGGAGACCAACCCCGCCCAAGCAGGAGGCCCAAGGACCGGAGCCCTGACCTCCCTGCGCTCCCTGTTCCGCTCCGACCCCTGTCCCA  
 GTAA

124/178

Fig. 81A

## 22. 2003 CON 10 CD gag . PEP

MGARASVL<sup>SG</sup>CKL<sup>DE</sup>WEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLLETSEGCKQII<sup>GO</sup>LQPAIQ<sup>TS</sup>EGSEIKSLYNTVATLYCVHERI  
 KVTDTKEALDKIEEEQTKSKKKAQ<sup>AT</sup>ADTGNSSQVSQNYPIVQNLQGMVHQP<sup>LS</sup>PRTLNAWVKVIEEKAFSPEVIMFSALSEGATPQDL  
 NTMLNTVGGHQAAMQMLKETINEEAAEWDR<sup>LP</sup>HPVQAGPVAPGQIREPRGSDIAGTTSTLQEQIRWMTSNPPIPVGEIYKRWIILGLNKIVRM  
 YSPVSILDIRQGPKEPERDYVDRFYKTLRAEQASQDVKNWMTETLLVQ<sup>AN</sup>PDCKTILKALGPAATLEEMMTACQGVGGPSHKARVLA<sup>EAM</sup>S  
 QATSGNAIMMORGNFKGPKKI<sup>IK</sup>CNFCGKEGHI<sup>AK</sup>NCRAPRKKGCKWKCGR<sup>EG</sup>HQMKDCTERQANFLGKIWPSNKG<sup>RP</sup>GNFLQSRPEPTAPPA  
 ESFGFGEIITPSQKEQKDKELHPLASLKS<sup>LF</sup>GN<sup>DP</sup>LSQ\$

Fig. 81B

## 2003 CON 10 CD gag . OPT

ATGGCGC<sup>CG</sup>CGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGAGTGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGTACCGCCT  
 GAAGCACCTGGTGTGGCCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCCGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCG  
 GCCAGCTGACGCCGCCATCCAGACCGGCTCCGAGGAGATCAAGTCCCTGTACAACACCGTGGCCACCCCTGTACTGCGTGACGAGCGCATC  
 AAGGTGACCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGACCAAGTCCAAGAAGAGGCCAGCAGGCCACCGCCGACACCGG  
 CAACTCCTCCAGGTGTCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACACCCCTGTCCCCCGCACCCCTGAACG  
 CCTGGGTGAAGGTGATCGAGGAGAAGGCCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCAGGACCTG  
 AACACCATGCTGAACACCGTGGCGGCCACAGGCCGCCATGAGATGCTGAAGGAGACCATCAACGAGGAGCGCGCGAGTGGGACCGCCT  
 GCACCCCGTGCAGGCCGCCCGTGGCCCCCGGCCAGATCCGGAGCCCCCGCGCTCCGACATCGCCGGCACCACTCCACCCCTGCAGGAGC  
 AGATCCGCTGGATGACCTCCAAACCCCCCATCCCGTGGCGGAGATCTACAAGCGCTGGATCATCTGGGCTGAACAAGATCGTGCGCATG  
 TACTCCCCCGTGTCCATCTGGACATCCGCCAGGGCCCCAAGGAGCCCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACCCCTGCGCGCCGA  
 GCAGGCTCCCGAGGACGTGAAGAACTGGATGACCGAGACCCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG  
 GCCCGCGCCACCCCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCCCTCCCAACAAGGCCCGCGTGTGGCCGAGGCCATGTCC  
 CAGGCCACCTCCGGCAACGCCATCATGATGACGCGGCAACTTCAAGGGCCCCAAGAAGATCATCAAGTGTCTCAACTGCGGCAAGGAGGG  
 CCACATCGCCCAAGAACTGCCGCGCCCCCGCAAGAAGGCTGCTGGAAGTGGCGCGGAGGCCACCCAGATGAAGGACTGCACCGAGCGCC  
 AGGCCAACTTCTGGGCAAGATCTGGCCCTCCAAACAAGGGCGCCCCCGGCAACTTCTGCAGTCCGCCCCGAGCCCCACCGCCCCCCCCC  
 GAGTCTTCCGCTTCGGCGAGGAGATCACCCCTCCCAAGAAGCAGGAGCAAGGAGCTGCACCCCTGGCCTCCCTGAAAGTCCCT  
 GTTCGGCAACGACCCCTGTCCCCAGTAA



Fig. 83A

24. 2003 CON 12 BF.gag.PEP  
 MGARASVL<sup>SG</sup>GEID<sup>RW</sup>EKIRLRPGGKKYRLKHIVWASRELERFAVNPGLLETSEGCRKIIGQLQPSLQTGSEELRSLYNTIAVL<sup>YF</sup>VHQKV  
 EVKDTKEALDKLEEEQNK<sup>SQ</sup>KTQQAADKGV<sup>SQ</sup>NYPIVQNLOGMVHQALSPRTLN<sup>AW</sup>KVVEEKAFSP<sup>EVI</sup>PMFSALSEGATPQDLN<sup>TML</sup>  
 NTVGGHQAAMQLKDTINEEAAEWDRLHPVHAGPIPPGQMRPRGSDIAGTTSTLQEQIQWMTSNPPVPVGEIYKRWIILGLNKIVRMYS<sup>SPV</sup>  
 SILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQ<sup>AN</sup>PDCKTILKALGPGATLEEMMTACQGVGGPGHKARVLAEAMSQV<sup>TN</sup>  
 TTVMQKSNEFKGORRIVKCFNCGKEGHIAKNCRAPRKKGCKGREGHQMKDCTERQANFLGKIWPSNKGKRPGNFLQNRPEPTAPPAES<sup>FGF</sup>  
 GEEITSPKQEQKDEGLYPPLASLKLFGNDP\$

Fig. 83B

2003 CON 12 BF.gag.OPT  
 ATGGCGCCCGGCTCCGTGCTGTCCGGCGCGGAGCTGGACCGTGGGAGAAGATCCGCTGCGCCCCGCGGCGGAAGAAGTACCGCCT  
 GAAGCACATCGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCGTGAACCCGCGCTGTGGAGACCTCCGAGGGCTGCCGCAAGATCATCG  
 GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCATCGCCGTGTACTTCTGTGCACCAAGAGTG  
 GAGTGAAGGACACCAAGGAGGCCCTGGACAAGCTGGAGGAGGAGCAGACAAGTCCACAGAGAAGACCCAGAGGCCGCCGCCGACAAAGG  
 CGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACAGGCCCTGTCCCCCGCACCTGAACGCCCTGGGTGAAGG  
 TGGTGGAGGAGAAGCCCTTCTCCCCAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCGCAGGACCTGAACACCATGCTG  
 AACACCGTGGCGGCGCACAGGCCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCCGCGAGTGGACCGCTGCACCCCGCTGCA  
 CGCCGGCCCCATCCCCCGGCGAGATGCGCGGAGCCCCGGCTCCGACATCGCCGCGCACCATCCACCCCTGCAGGAGCAGATCCAGTGGA  
 TGACCTCCAAACCCCCCGTGCCTGGCGGAGATCTACAAGCGCTGGATCATCTGGGCCGTGAACAAGATCGTGGCATGTACTCCCCCGTG  
 TCCATCTGGACATCCGCGAGGCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCTTGGCGCCGAGCAGGCCACCCA  
 GGAGGTGAAGGGCTGGATGACCGACACCTGTGTGTGAGACGCCAACCCCGACTGCAAGACCATCTGAAGGCCCTGGGCCCGCGGCCA  
 CCTGGAGGAGATGATGACCGCTGCCAGGGCGTGGCGGCGCCCGGCAAGGCCCGGTGCTGGCCGAGGCCATGTCCAGGTGACCAAC  
 ACCACCGTATGATGCAGAAGTCCAACTCAAGGGCCAGCGCGCATCGTGAAGTCTTCACTGCGGCAAGGAGGCCACATCGCCCAAGAA  
 CTGCGCGCCCCCGCAAGAAGGGCTGTGGAAGTGGCGCGGAGGGCCACAGATGAAGGACTGCACCCGAGCGCCAGGCCAACTTCCTGG  
 GCAAGATCTGGCCCTCCAAACAAGGGCGGCCCGGCAACTTCCTGCAGAAACCGCCCCGAGCCACCGCCCCCGGAGTCTTCGGCTTC  
 GCGGAGGAGATCACCCCTCCCCCAAGCAGGAGCAGAGGAGGCGCTGTACCCCCCTGGCCTCCCTGAAGTCCCTGTTCGGCAACGA  
 CCCCTAA

126/178

Fig. 84A

25. 2003 CON 14 BG gag.PEP  
 MGARASVL<sup>~</sup>SGGK<sup>~</sup>DAWEKIRLRPGGKKYRMKHLVWASRELERFALNPDLL<sup>~</sup>ETAEGCQQIMQLQ<sup>~</sup>PALQ<sup>~</sup>TGTEEIRSLFNTVATLYCVHQKI  
 EVKDTKEALEEVEKAQKKSQKKQQAAMDEGNN<sup>~</sup>SQASQNYPIVQNAQGMVHQ<sup>~</sup>AI<sup>~</sup>SPRTLN<sup>~</sup>AWKVVEEKA<sup>~</sup>FSPEV<sup>~</sup>IPMFSALSEGATPQDIN  
 TMLNTVGGHQAA<sup>~</sup>MQMLKDTINEEAAEWD<sup>~</sup>RMHPQ<sup>~</sup>QAGPIPPGQIREPRGSDIAGTTSTLQEQIRWMTSNPP<sup>~</sup>IPVGEIYKRWI<sup>~</sup>ILGINKIVRMV  
 SPVSILDIRQGPKEPF<sup>~</sup>RDYVDRFFKTLRAEQATQEVKGWMTD<sup>~</sup>TLVQ<sup>~</sup>NANPDCKTILRALPGATLEEMMTACQGVGSPSHKARVLAEAMSQ  
 ASGATIMMQKSNFKGPRRN<sup>~</sup>IKCFNCGKEGHLARNCRAPRKKGCWKCKGKEGHQMKDCTESKANFLGKIWPSNKG<sup>~</sup>RPGN<sup>~</sup>FLQNRPEPTAPPAES  
 FGFGEIAPSPKQEPKEKEIYPLASLSLFGSDP\$SQ\$

Fig. 84B

2003 CON 14 BG gag.OPT  
 ATGGGCGC<sup>~</sup>CGGCTCCGTGCTGTCCGGGGCAAGCTGGACGCCTGGGAGAAGATCCGCCCTGGCCCCGGCGGCAAGAAGTACCGCAT  
 GAAGCACTGGTGTGGGCTCCCGGAGCTGGAGCGCTTCGCCCTGAACCCGACCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG  
 GCCAGCTGACGCCGCCCTGCAGACCGGACCGAGGAGATCCGCTCCCTGTTCAACACCGTGGCCACCTGTACTGCGTGCA<sup>~</sup>CCAGAA<sup>~</sup>GATC  
 GAGTGAAGGACACCAAGGAGGCCCTGGAGGAGGTGGAGAGGCCCAAGAGTCCCAAGAGAAGCAGCAGGCGCCCATGGACGAGGGCAA  
 CAATCCCAGGCTCCAGAACTACCCATCGTGCAGAACGCCCAAGGGCCAGATGGTGCAACAGGCCATCTCCCCCGCACCCCTGAACGCCT  
 GGTGAAGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC<sup>~</sup>CCAGGACCTGAAC  
 ACCATGCTGAACACCGTGGCGGCCACCAAGCGGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGCGCCGCGAGTGGGACCGCATGCA  
 CCCCAGAGCGCGGCCCATCCCCCGGCGAGATCCGCGAGCCCGCGGCTCCGACATCGCCGGCACCACTCCACCTGCAGGAGCAGA  
 TCCGCTGGATGACCTCCAACCCCCCATCCCCGTGGCGGAGATCTACAAGCGTGGATCATCTTGGGCCCTGAACAAGATCGTGGCATGTAC  
 TCCCCGTGTCCATCTGTGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTCAAGACCTTGGCGGCCGAGCA  
 GGCAACCCAGAGGTGAAGGCTGGATGACCGCATCCAGGCGCTGGCGGCCCTCCCA<sup>~</sup>CAAGGCCCGCTGTGGCCGAGGCCATCTCGGCCCTGGGCC  
 CCGCGCCACCTGGAGGAGATGATGACCGCTGCCAGGCGCTGGCGGCCCTCCCA<sup>~</sup>CAAGGCCCGCTGTGGCCGAGGCCATGTCCCAG  
 GCCTCCGGCGCCACCATCATGATGCAGAA<sup>~</sup>GTCCA<sup>~</sup>ACTCAAGGGCCCCCGCGCAACATCAAGTGTTC<sup>~</sup>CAACTGGGCAAGGAGGCCACCT  
 GGCCGCAACTGCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCCACCA<sup>~</sup>GATGAAGGACTGCA<sup>~</sup>CCGAGTCCAAGGCCA  
 ACTTCTGGGCAAGATCTGGCCCTCCAACAAGGGCGGCCCGCGCAACTTCTGCAGAACCGCCCCGAGCCACCGCCCCCGCGAGTCC  
 TTCGGCTTCGGCGAGGAGATCGCCCCCTTCCCCCAAGCAGGAGCCCAAGGAGAAGGAGATCTACCCCTTGGCCTTCCCTGAAGTCCCTGTTCGG  
 CTCCGACCCCTAATCCCAGTAA

127/178



128/178

Fig. 85A

## 31. 2003 CONS nef.PEP

MGGKWSKSSIVGWPAVRERIRRTPPAAEGVGAVSQDLDKHGAISSNTAATNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKGAFDLSHFLK  
 EKGGLDGLIYSKKRQEIILDLWVYHTQGYFPDQNYTPGIRYPLTFGWCFKLVDPDPEEVEEANEENNCLLHPMCQHGMEDEREVLMMWK  
 FDSRLALRHIARELHPEFYKDC\$

Fig. 85B

## 2003 CONS nef.OPT

ATGGCGGCAAGTGGTCCAAGTCCATCGTGGGCTGGCCCGCGTGGCGGAGCGCATCCGCCCGCACCCCCCGCGGAGGGCGGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCAGCGGCCATCACCTCTCAACACCGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGTGGCTTCCCGTGGCGCCCGAGGTGCCCTGCGCCCATGACCTACAGGGCGCCTTGGACCTGTCCCACTTCCCTGAAG  
 GAGAAGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCGCCCGCATCGCTACCCCTGACCTTCCGCTGGTGTCAAGCTGGTGCCCGTGACCCCGAGGAGGTGG  
 AGGAGGCCAACGAGGGCGAGAACAACTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACGCGAGGTGCTGATGTGGAAG  
 TTGCACTCCCGCCTGGCCCTGGCCACATCGCCCGGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 86A

## 32. 2003 M. GROUP.anc nef.PEP

MGGKWSKSSIVGWPAVRERIRRTAPAAEGVGAVSQDLDKHGAISSNTAATNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKAAFDLSHFLK  
 EKGGLDGLIYSKKRQEIILDLWVYHTQGYFPDQNYTPGIRYPLTFGWCFKLVDPDPEEVEEANEENNCLLHPMCQHGMEDEREVLMMWK  
 FDSRLALRHIARELHPEFYKDC\$

Fig. 86B

## 2003 M GROUP.anc nef.OPT

ATGGCGGCAAGTGGTCCAAGTCCATCGTGGGCTGGCCCGCGTGGCGGAGCGCATGCCGCCGACCGCCCCCGCGGAGGGCGGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCAGCGGCCATCACCTCTCAACACCGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCCAGG  
 AGGAGGAGGAGTGGCTTCCCGTGGCGCCCGAGGTGCCCTGCGCCCATGACCTACAGGGCGCCTTGGACCTGTCCCACTTCCCTGAAG  
 GAGAAGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCGCCCGCATCCGCTACCCCTGACCTTCCGCTGGTGTCAAGCTGGTGCCCGTGACCCCGAGGAGGTGG  
 AGGAGGCCAACGAGGGCGAGAACAACTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACGCGAGGTGCTGATGTGGAAG  
 TTGCACTCCCGCCTGGCCCTGGCCACATCGCCCGGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 87A

## 33. 2003 CON A nef .PEP

MGKWSKSSIVGWPDIRIRRTPPAAKGVAVSQDLDKYGAVTINNTAATQASCAWLEAQEEEEVEGFPVRPQVPLRPMTFKGAFDLSFFL  
KEKGLDGLIYSQKRQEIILDLWYNTQGYFPDWNQYTPGPTRFPLTFGWCFKLVVDPEDEVEEATEGENNCLLHPICQHGMDDEEKEVLMW  
KFDSRLARRHIALEMHPFYKDC\$

Fig. 87B

## 2003 CON A nef .OPT

ATGGCGGCAAGTGGTCCAAAGTCTCCATCGTGGGTGGCCCGACATCCGCGAGCGCATCCGCGGCACCCCGCCGCAAGGCGTGGG  
CGCGTGTCCAGGACCTGGACAAGTACGGCGCGGTGACCATCAACAACACCGCCGCCACCCAGGCTCCTGCGCCTGGCTGGAGGCCCAGG  
AGGAGGAGGAGGAGGTGGCTTCCCGTGGCGCCCGCAGGTGCGCCCATGACCTTCAAGGGCGCCTTCGACCTGTCTTCTTCTG  
AAGAGAAGGGCGGCTGGACGGCCTGATCTCCAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGCTACTTCCC  
CGACTGGCAGAACTACACCCCGCGCCCGGCAACCGCTTCCCGTGGCTTGGCTTCAAGCTGGTGGCGTGGACCCCGACGAGG  
TGGAGGAGGCCACCGAGGGCGAGAACAACTGCTGTGACCCCATCTGGCAGCACGGCATGGACGACGAGGAGAAGGAGGTGCTGATGTGG  
AAGTTCGACTCCCGCCTGGCCCGCCGACATCGCCCTGGAGATGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 88A

## 34. 2003 CON A1 nef .PEP

MGKWSKSSIVGWPPEVRERMRRTPPAATGVGAVSQDLDKHGAVTSSNINHPSCVWLEAQEEEEVEGFPVRPQVPLRPMTYKGALDLSHFLKEK  
GGLDGLIYSRKRQEIILDLWYHTQGYFPDWNQYTPGPGIRYPLTFGWCFKLVVDPEVEKATEGENNCLLHPICQHGMDDEEREVLKWKFD  
SRLALKHRAQELHPEFYKDC\$

Fig. 88B

## 2003 CON A1 nef .OPT

ATGGCGGCAAGTGGTCCAAAGTCTCCATCGTGGGTGGCCCGAGGTGCGGAGCGCATCGCGCGCACCCCGCCCGCCGCAAGGCGTGGG  
CGCGTGTCCAGGACCTGGACAAGCACGGCGCGGTGACCTCTCCAACATCAACACCCCTCCTGCGTGTGGCTGGAGGCCAGGAGGAGG  
AGGAGTGGGTTCCTCGTGGCGCCCGCAGGTGCGCCCATGACCTACAAAGGGCGCCTGGACCTGTCCCACTTCTGAAGGAGAAG  
GGCGCCTGGACGGCCTGATCTACTCCGCAAGCGCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGCTACTTCCCCGACTGGCA  
GAACTACACCCCGCGCCCGCATCCGCTACCCCTGACCTTCGGCTGGTCTTCAAGCTGGTGGCGTGGACCCCGACGAGGTGGAGAAGG  
CCACCGAGGGCGAGAACAACTCCCTGCTGACACCCCATCTGCCAGCACGGCATGGACGACGAGGAGCGGAGGTGCTGAAGTGAAGTTCGAC  
TCCCGCCTGGCCCTGAAGCACCGCGGCCAGGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

129/178

130/178

Fig. 88C

## 35. 2003 A1.anc nef.PEP

MGGKWSKSSIVGWPEVRERMRRTPPAAKGVAVSQDLDKHGAVTSSNTAANNPGCAWLEAQEEEEVGFPVRPQVPLRPMYKGAFDLSHFLK  
 EKGGLDGLIYSKKRQEIILDLVYHTQGYFPDQWNYTPGPGIRYPLTFGWCFKLVPVDPAPAEVEEATEGENNSLLHPICQHGMDDEREVLMWK  
 FDSRLALKHRARELHPEFYKDC\$

Fig. 88D

## 2003 A1.anc nef.OPT

ATGGCGGCAAGTGGTCCAAAGTCTCCATCGTGGGCTGGCCCGAGGTGCGGAGCGCATGCGCGCACCCCCCGCCGCAAGGGCGGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCAGGCGCCGTGACCTCTCCAAACACGCGCCCAACAACCCGGCTGCGCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCGCCAGGTGCCCTCGGCCCATGACCTACAAAGGCGCCTTCGACCTGTCCACTTCTCTGAAG  
 GAGAAGGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCCGGCCCGCATCCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGTGTCCTGGACCCCGCCGAGGTGG  
 AGGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGACGAGGAGCGGAGGTGCTGATGTGGAAG  
 TTCGACTCCCGCTGGCCCTGAAGCACCGCGCGGAGCTGCACCCGAGTCTACAAGGACTGCTAA

Fig. 89A

## 36. 2003 CON A2 nef.PEP

MGGKWSKSSIVGWPAIRERMRKRTPPAAEGVAVSQDLATRGAVTSSNTAATNPDCAWLEAQEEEEVGFPVRPQVPLRPMYKGAFDLSHFL  
 KEKGGLDGLIYSQKRQDILDLVYHTQGYFPDQWNYTPGPGTRYPLTFGWCFKLVPVDPSEVEEATEGENNSLLHPICQHGIEDPEREVLRW  
 KFDSRLALRHRARELHPEFYKDC\$

Fig. 89B

## 2003 CON A2 nef.OPT

ATGGCGGCAAGTGGTCCAAAGTCTCCATCGTGGGCTGGCCCGCATCCGCGAGGCGATGCGCAAGCGCACCCCCCGCCGCGAGGGCGGT  
 GGGCGCGGTGTCCAGGACCTGGCCACCCGGCGCCGTGACCTCTCCAACACGCGCCCAACCAACCCGACTGCGCTGGCTGGAGGCC  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCCCGCCAGGTGCCCTGCGCCCATGACCTCAAGGGCGCCTTCGACCTGTCCACTTCCTG  
 AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCCAGAAGCGCCAGGACATCTGGACCTGTGGGTGTACACACCCAGGGCTACTTCCC  
 CGACTGGCAGAACTACACCCCGGCGCCGCAACCCCTGACCTTCGGTGTGCTTCAAGCTGGTGGCCGTGGACCCCTCCGAGG  
 TGGAGGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATCGAGGACCCCGAGGCGGAGGTGCTGGCGCTGG  
 AAGTTCGACTCCCGCTGGCCCTGGCGCCACCGGGCCCGGAGCTGCACCCCGAGTCTCTACAAGGACTGCTAA

131/178

*Fig. 90A***37. 2003\_CON\_B nef.PEP**

MGGKWSKRSVVGWPTVRRMRRAEPAADGVGAVSRDLEKHGAITSSNTAANNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKGAIDLSHFLK  
 EKGGLEGLIYSQKRQDILDLWVYHTQGYFPDWQNYTPGPIRYPLTFGWCFKLVPVEPEKVEEANEGENNSLLHPMSLHGMDDPEREVLVWK  
 FDSRLAFHHMARELHPEYKDC\$

*Fig. 90B***2003\_CON-B nef.OPT**

ATGGGCGGCAAGTGGTCCAAAGCGCTCCGTGGTGGGCTGGCCACCGTGCGGAGCGCATGCGCGCGCGAGCCCCCGCCGACGGCGTGGG  
 CGCCGTGTCCCGGACCTGGAGAGCACGGCGCCATCACCTCCTCCAACACCGCCGCCAACAAACGCGGACTGCGCCTGGCTGGAGGCCACAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCAAGGTGCCCTGGCGCCCATGACCTACAAGGGCGCCCTGGACCTGTCCCACTTCCTGAAG  
 GAGAAAGGCGGCTGGAGGGCTGATCTACTCCAGAAAGCGCCAGGACATCTGGACCTGTGGGTGTACCAACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCCCGGCATCCGTACCCCTGACCTTCGGTGGTCAAGCTGGTGGCGGAGCCCCGAGAGGTTGG  
 AGGAGGCCAACGAGGCGAGAACAACTCCCTGCTGCACCCCATGTCCCTGCACGGCATGGACGACCCCGAGCGGAGGTGCTGTGGAAG  
 TTCGACTCCCGCCTGGCCTTCCACCACATGGCCCCGGAGCTGCACCCGAGTACTACAAGGACTGCTAA

*Fig. 90C***38. 2003\_B.anc nef.PEP**

MGGKWSKSSMGGWPVRRMRRAEPAADGVGAVSRDLEKHGAITSSNTAATNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKAALDLSHFLK  
 EKGGLEGLIYSQKRQDILDLWVYHTQGYFPDWQNYTPGPIRYPLTFGWCFKLVPVEPEKVEEATEGENNSLLHPMCQHGMDDPKEVLVWK  
 FDSRLAFHHMARELHPEYKDC\$

*Fig. 90D***2003\_B.anc nef.OPT**

ATGGGCGGCAAGTGGTCCAAAGTCCATGGGCGGCTGGCCCGCGCTGGCGGAGCGCATGAAGCGCGCGCGAGCCCCCGCCGACGGCGTGGG  
 CGCCGTGTCCCGGACCTGGAGAGCACGGCGCCATCACCTCCTCAACACCGCCGCCAACACCGGACTGCGCCTGGCTGGAGGCCACAGG  
 AGGAGGAGGAGGTGGGCTTCCCGTGGCGCCCAAGGTGCCCTGGCGCCCATGACCTACAAGCGCCCGCTGGACCTGTCCCACTTCCTGAAG  
 GAGAAAGGCGGCTGGAGGGCTGATCTACTCCAGAAAGCGCCAGGACATCTGGACCTGTGGGTGTACCAACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCCCGGCATCCGTACCCCTGACCTTCGGTGGTCAAGCTGGTGGCGGAGCCCCGAGAGGTTGG  
 AGGAGGCCAACGAGGCGAGAACAACTCCCTGCTGCACCCCATGTGCCAGCACGGCATGGACGACCCCGAGAGGAGGTGCTGTGGAAG  
 TTCGACTCCCGCCTGGCCTTCCACCACATGGCCCCGGAGCTGCACCCGAGTACTACAAGGACTGCTAA

132/178

Fig. 91A

39. 2003 CON 02 AG nef. PEP  
 MGKWSKSSIVGWPVKVRERIRQTTPAATGVGAASQDLDRHGAI TSSNTAATNADCAWLEAQEEEEVEGFPVRPQVPLRPMTYKAAVDLSHFLK  
 EKGGLEGLIYSKKRQEI LLDLVYHTQGFDPDWNQNTPGPTRFPLTFGWCFKLVPMDPAEVEEANEGENNSLLHPICQHGMEDEDEDREVLVWR  
 FDSLSAFKHRARELHPEFYKDC\$

Fig. 91B

2003 CON 02 AG nef. OPT  
 ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGGTGGCCCAAGGTGGCGAGCGCATCCGCCAGACCCCGCCCGCCACCCGGCGTGGG  
 CGCCGCTCCAGGACCTGGACCGCCACGCGCCATCACCTCTCAACACCGCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGTGGCTTCCCGTGGCGCCCGAGGTGCCCTGGCCCATGACCTACAAGGCCGCGTGGAACCTGTCCCACTTCTTCCCGA  
 GAGAAGGCGGCTGGAGGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCAACCCAGGGCTTCTTCCCGA  
 CTGGCAGAACTACACCCCGGCCCCGACCCGCTTCCCGTGGCTGGTGTCAAGCTGTGGTGTGACCCCGCCGAGGTGGTGGCGG  
 AGGAGGCCAACGAGGCGAGAACACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGGACCGGAGGTGCTGGTGGCGG  
 TTGCACTCCTCCCTGGCCTTCAAGCACCGCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 92A

40. 2003 CON C nef. PEP  
 MGKWSKSSIVGWPVAVRERIRRTEPAEGVGAASQDLDRHGALTSSNTATNADCAWLEAQEEEEVEGFPVRPQVPLRPMTYKAAFDLSFFL  
 KEKGGLEGLIYSKKRQEI LLDLVYHTQGYFPDWNQNTPGPGVRYPLTFGWCFKLVDPDPREVEEANEGENNCLLHPMSQHGMEDEDEDREVLKW  
 KFDShLARRHMARELHPEYKDC\$

Fig. 92B

2003 CON C nef. OPT  
 ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGGTGGCCCGCGTGGCGAGCGCATCCGCCGACCCGAGCCCGCCCGCGGCGTGGG  
 CGCCGCTCCAGGACCTGGACAGCACGGCCCTGACCTCTCAACACCGCCACCAACAGCCGACTGCGCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGTGGCTTCCCGTGGCGCCCGAGGTGCCCTGGCCCATGACCTACAAGGCCGCTTGAACCTGTCTTCTTCTG  
 AAGGAGAAGGCGGCTGGAGGCTGATCTACTCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCAACCCAGGCTACTTCCC  
 CGACTGGCAGAACTACACCCCGGCCCCGCTGACCTTCCGCTGGTGTCAAGCTGGTGGACCCCGTGGACCCCGGAGG  
 TGGAGGAGGCCAACGAGGCGGAGAACAACTGCCTGTGCACCCCATGTCCAGCACGGCATGGAGGACGAGGAGGTGCTGAAGTGG  
 AAGTTCGACTCCCACTGGCCCGCCGACATGGCCCGGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

133/178

Fig. 92C

## 41. 2003 C.anc nef.PEP

MGKWSKSSIVGWPAVRERMRRTPEAAEGVGAASQDLDKHGALTSSNTAANNADCAWLEAQEEEEVEVFPVRPQVPLRPMTYKAAFDLSFFL  
KEKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWNQYTPGPGVRYPLTFGWCFKLVDPREVEEANEENNCILLHPMSQHGMEDDEDREV LKW  
KFDSHLARRHMARELHPEYKDC\$

Fig. 92D

## 2003 C.anc nef.OPT

ATGGGGCAAGTGGTCCAAAGTCTCCATCGTGGGTGGCCCGCGGTGGCGAGCGCATGCGCCGCAACGAGCCCCCGCGGAGGGCGTGGG  
CGCCGCTCCAGGACCTGGACAAGCACGGGCCCTGACCTCTCAACACCGCCGCCAACACGCCGACTGGCCTGGCTGGAGGCCCAGG  
AGGAGGAGGAGGAGGTGGCTTCCCCGTGGCCCCCAGGTGCCCTGCGCCCATGACCTACAGGCCGCTTCGACCTGTCTTCTCTG  
AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCAAGAAGCGCCAGAGATCCTGGACCTGTGGGTGTACCAACCCAGGGCTACTTCCC  
CGACTGGCAGAACTACACCCCCGGCCCGGTGCGCTACCCCTGACCTTCGGCTGGTGTCAAGTGGTGCCGTGGACCCCGCGAGG  
TGGAGGAGGCCAACGAGGGCGAGAACAACTGCTGTGACCCCATGTCCACGACCGCATGGAGACGAGGACCGGAGGTGCTGAAGTGG  
AAGTTCGACTCCCACTGGCCCGCGGACATGGCCCGGAGTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 93A

## 42. 2003 CON D nef.PEP

MGKWSKSSIVGWPAIRERIRRTPEAADGVGAVSRDLEKHGALTSSNTAATNADCAWLEAQEEDEEVFPVRPQVPLRPMTYKAAALDLSHFL  
KEKGGLEGLVWSQKRQEILDLWVYNTQGFDPWNQYTPGPGIRYPLTFGWCFELVPDPREEVEATEGENNCILLHPMCQHGMEDPEREVLMMW  
RFNSRLAFEHKARVLHPEFYKDC\$

Fig. 93B

## 2003 CON D nef.OPT

ATGGGGCAAGTGGTCCAAAGTCTCCATCGTGGGTGGCCCGCCATCCGGAGCGCATCCGGCGACCGAGCCCCCGCGGAGGGCGTGGG  
CGCCGTGTCCCGGACCTGGAGAAGCACGGGCCATCACCTCTCAACACCGCCGCCAACAGCCGACTGGCCTGGCTGGAGGCCCAGG  
AGGAGGACGAGGAGGTGGCTTCCCCGTGGCCCCCAGGTGCCCTGCGCCCCATGACCTACAAAGCCGCCCTGGACCTGTCCCACTTCTG  
AAGGAGAAGGGCGGCTGGAGGGCTGGTGTGGTCCAGAGCGCCAGGATCCTGGACCTGTGGGTGTACAAACCCAGGGCTTCTTCCC  
CGACTGGCAGAACTACACCCCGCGGCGCATCCGCTACCCCTGACCTTCGGCTGGTGTTCGAGCTGGTGGACCCCGAGGAGG  
TGGAGGAGGCCACCGAGGGCGAGAACAACTGCCTGTGCACCCCATGTGCCAGCACGGCATGGAGGACCCCGAGCGGAGGTGCTGATGTGG  
CGCTTCAACTCCCGCTGGCCTTCGAGCACAAAGGCCCGCGTGTGCACCCCGAGTTCTACAAGGACTGCTAA

134/178

Fig. 94A

43. 2003 CON F1 nef. PEP  
 MGKWSKSSIVG<sup>W</sup>PAVRERMRPTPPAAEGVGAVSQDLERRGAI<sup>T</sup>SSNTGATNPDLAWLEAQEEEEVGFPVRPQVPLRPM<sup>T</sup>YKGAVDLSHFLK  
 EKGGLEGLIYSKKRQEI<sup>L</sup>DLWVYHTQGYFPDWQNYTPGPIRYPLTFGWCFKLV<sup>P</sup>VDPEEVEKANE<sup>G</sup>ENNC<sup>L</sup>LHPMSQHMEDEDE<sup>R</sup>EVLIWK  
 FDSRLALR<sup>H</sup>RIARERHPEFYQD\$

Fig. 94B

2003 CON F1 nef. OPT  
 ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGGTGGCCCGCGGTGGCGAGCGCATGGCCCCACCCCCCGCGCGAGGCGGTGGG  
 CGCGGTGTCCCAGGACCTGGAGCGCGCGCCATACCTCTCCAACACGCGGCCACCAACCCGACCTGGCCCTGGCTGGAGGCCCAGG  
 AGGAGGAGGAGTGGGTTC<sup>C</sup>CGGTGGCGCCCGCAGGTGCCCTGGCGCCCATGACCTACAAAGGCGCGCTGGACCTGTCCCAC<sup>T</sup>TCCCTGAAG  
 GAGAAGGCGGCGCTGGAGGCGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCAACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCGCATCGCTACCCCTGACCTTCGGCTGGTCTCAAGCTGGTGGCCCGTGGACCCCGAGGAGGTGG  
 AGAAGGCCAACGAGGGCGAGAACAACTGCCGTGTGCACCCCATGTCCCAGCACCGCATGGAGGACCGGAGGCTGCTGATCTGGAAG  
 TTCGACTCCCCGCTGGCCCTGGCGCCACATCGCCCCGCGAGCGCCACCCCGAGTCTACAGGACTAA

Fig. 95A

44. 2003 CON F2 nef. PEP  
 MGKWSKSSIVG<sup>W</sup>PTIRERIRRT<sup>P</sup>VAAEGVGAVSQDLKKGAI<sup>T</sup>SSNTRATNADLAWLEAQEEDEEVGFVRPQVPLRPM<sup>T</sup>YKAAFDLSHFLK  
 EKGGLEGLIYSKKRQEI<sup>L</sup>DLWVYHTQGYFPDWQNYTPGPGTRYPLTFGWCFKLV<sup>P</sup>VDPEEVEKANE<sup>G</sup>ENNC<sup>L</sup>LHPMSLHGMEDEDE<sup>R</sup>EV<sup>L</sup>KKWK  
 FDSRLALR<sup>H</sup>RIARERHPEYYKD\$

Fig. 95B

2003 CON F2 nef. OPT  
 ATGGCGGCAAGTGGTCCAAGTCTCCATCGTGGGTGGCCCGCACCATCCGCGAGCGCATCCGCCGACCCCCGTGGCCCGAGGCGGTGGG  
 CGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCTCCAACACCGCGCCACCAACGCCGACCTGGCCCTGGCTGGAGGCCCAGG  
 AGGACGAGGAGGTGGGTTC<sup>C</sup>CGGTGGCGCCCGCAGGTGCCCTGGCGCCCATGACCTACAAAGCGCGCTTCGACCTGTCCCAC<sup>T</sup>TCCCTGAAG  
 GAGAAGGCGGCGCTGGAGGCGCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCAACACCCAGGGCTACTTCCCCGA  
 CTGGCAGAACTACACCCCGCGCACCCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGGCCCGTGGACCCCGAGGAGGTGG  
 AGAAGGCCAACGAGGGCGAGAACAACTGCCGTGTGCACCCCATGTCCCTGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTGGGAAG  
 TTCGACTCCCCGCTGGCCCTGGCGCCACATCGCCCCGCGAGCGCCACCCCGAGTACTACAAGGACTAA

Fig. 96A

## 45. 2003 CON G nef.PEP

MGKWSKSSIVGWP<sup>EV</sup>RRIRIQTPPAEGVGAVSQDLARHGAI<sup>TS</sup>NTAANNPDCAWLEAQEEDSEVGFVPRPQVPLRPM<sup>TY</sup>KGAFDLSFFL  
KEKGGLDGLIYSKKRQDILD<sup>WV</sup>YNTQGFEPDWQNYTPGPGTRFPLTFGWCFKLVPM<sup>DP</sup>AEVEEANKGENNSLLHPICQHGMED<sup>ED</sup>REVLVW  
RFDSSLARRHIARELHPEYKDC\$

Fig. 96B

## 2003 CON G nef.OPT

ATGGCGGCAAGTGGTCCCAAGTCTCCATCGTGGGCTGGCCCGAGGTGGCGGAGCGCATCCGCCAGACCCCCCGCCCGCCGAGGGCGTGGG  
CGCCGTGTCCAGGACCTGGCCCGCCACGGCGCATCACCTCTCCAACACCGCGCCCAACAACCCGACTGCGCCTGGCTGGAGGCCCAGG  
AGGAGACTCCGAGGTGGGCTTCCCGTGGCCCGCCAGGTGCCCCGATGACCTACAAGGGCGCTTCGACCTGTCTTCTTCTG  
AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCCAAGAAGCGCCAGGACATCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCC  
CGACTGGCAGAACTACACCCCGGGCCCGGACCCGCTTCCCCCTGACCTTCGGCTGGTGTCAAGCTGGTGCCCATGGAACCCCGCCGAGG  
TGGAGGAGGCCAACAAAGGGCGAGAACAACTCCCTGCTGACCCCATCTGCCAGACGGCATGGAGGACGAGGACCGGAGGTGCTGCTGTGG  
CGCTTCGACTCTCCCTGGCCCGCCGACATCGCCCGGAGCTGCAACCCGAGTACTACAAGGACTGCTAA

135/178

Fig. 97A

## 46. 2003 CON H nef.PEP

MGKWSKSSIGGWP<sup>AI</sup>RRIRRAEPAAEGVAVSRDLRRGAVTINNTASTNPDSAWLEAQEEEEVEVGFVPRPQVPLRPM<sup>TY</sup>KGAFDLSHFL  
KEKGGLEGLIYSKKRQEILD<sup>WV</sup>YNTQGYFPDWQNYTPGPGERYPLTFGWCFKLV<sup>PD</sup>PQVEVEKANEGENNSLLHPICQHGMED<sup>EE</sup>REVLW  
KFDSRLAFRHHIARELHPEFYKDC\$

Fig. 97B

## 2003 CON H nef.OPT

ATGGCGGCAAGTGGTCCCAAGTCTCCATCGGCGGCTGGCCCGCCATCCGCCAGCGCATCCGCCGCGCCGAGCCCGCCGAGGGCGTGGG  
CGCCGTGTCCCGGACCTGGACCGCGCGGCGCGTGACCATCAACAACACCGCTCCACCAACCCGACTCCGCTGGCTGGAGGCCCAGG  
AGGAGGAGGAGGAGGTGGCTTCCCGTGGCCCGCCAGGTGCCCCATGACCTACAAGGGCGCTTCGACCTGTCCCACTTCTG  
AAGGAGAAGGGCGGCTGGAGGCTGATCTACTCCAAGAAGCGCCAGGAGATCTGGACCTGTGGGTGTACAACACCCAGGGCTACTCCC  
CGACTGGCAGAACTACACCCCGGGCCCGGAGCGCTACCCCTGACCTTCGGCTGGTGTCAAGCTGGTGCCCGTGAACCCCGAGG  
TGGAGAAGGCCAACGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGGACGAGGAGCGGAGGTGCTGATGTGG  
AAGTTCGACTCCCGCCTGGCTTCCGCCACATCGCCCGGAGCTGCACCCCGAGTCTACAAGGACTGCTAA



136/178

Fig. 98A

## 47. 2003 CON 01 AE nef .PEP

MGGKWSKSSIVGWPQVRERIKQTPPATEGVGAVSQDLDKHGAVTSSNMNADCVWLRQAEEEEVGFPVRPQVPLRPMTYKGAFDLSFFLKEK  
 GGLDGLIYSKKRQEIILDLWVYNTQGFPPDWQNYTPGPGIRYPLCFGWCFKLVPVDPREVEEDNKGENNCLLHPMSQHIGIEDEEREVLMMWKFD  
 SALARKHIARELHPEYKDC\$

Fig. 98B

## 2003 CON 01 AE nef .OPT

ATGGCGGCAAGTGGTCCAAAGTCCTCCATCGTGGGTGGCCCCAGGTGCGGAGCGCATCAAGCAGACCCCCCGCCACCGAGGGCGTGGG  
 CGCCGTGTCCAGGACCTGGACAAGCACGGCGCGTGACCTCCTCCAACATGAACAACGCCGACTGCGTGTGGCTGCGCGCCACAGGAGGAGG  
 AGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCCAGGTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCTTCTTCTGAAGGAGAAG  
 GCGGCGCTGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCCCGACTGGCA  
 GAACTACACCCCCCGGCATCCGCTACCCCCGTGGCTGCTCAAGCTGGTCCCCGTGGACCCCCCGGAGGTGGAGGAGG  
 ACAACAAGGGCGAGAACAACTGCCTGTGCACCCCCATGTCCAGCACGGCATCGAGGACGAGGAGCGCGAGGTGCTGATGTGGAAAGTTCGAC  
 TCCGCCCTGGCCCCGCAAGCACATCGCCCCGAGCTGCACCCCCAGTACTACAAGGACTGCTAA

Fig. 99A

## 48. 2003 CON 03 AE nef .PEP

MGGKWSKSSIVGWPQVRERIRRAPAPAARGVPVSQDLDKYGAVTSSNTAANNADCAWLEAQKEEEVGFPVRPQVPLRPMTYKGAFDLSHFL  
 KEKGLDGLIYSKKRQEIILDLWVYHTQGYFPDWQNYTPGPGIRFPPLTFGWYKLVVPVDPDEVEEATEGENNSLLHPICQHGMDDEEKEVLMW  
 KFD SRLALTHRARELHPEFYKDC\$

Fig. 99B

## 2003 CON 03 AE nef .OPT

ATGGCGGCAAGTGGTCCAAAGTCCTCCATCGTGGGTGGCCCCAGGTGCGGAGCGCATCCGCGCGCCCCCGCCCGCCCGCGCGGT  
 GGGCCCCGTGTCCAGGACCTGGACAAGTACGGCGCCGTGACCTCCTCAACACCGCCCAACACGCCGACTGCGCTGGCTGGAGGCCCC  
 AGAAGGAGGAGGAGTGGCTTCCCCGTGCGCCCCCAGGTGCCCCCAGTACCTACAGGGCGCCTTCGACCTGTCCCACTTCCTG  
 AAGAGAAGGGCGCCTGGACCGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCC  
 CGACTGGCAGAACTACACCCCCCGGCATCCGCTTCCCCCTGACCTTCGGCTGGTGTACAAGCTGGTGGCCGTGGACCCCCGACGAGG  
 TGGAGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCCATCTGCCAGCACGGCATGGACGACGAGGAGGAGGTGCTGATGTGG  
 AAGTTCGACTCCCGCTGGCCCTGACCCACCGCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA





140/178

Fig. 106A

## 55. 2003 CON 14 BG nef. PEP

MGKWSKCSIVGWEVRERIRRTPPAAVGVGAVSQDLAKHGAITSSNTAANNPDCAWLEAQEEDSEVGFVRPQVPLRPMYKGAFDLSFFL  
KEKGLDGLIYSKQRQDILDWVYNTQGFDPDQNYTPGPTRYPLTFGWCFKLEPVDPAEVEEATKGENNSLLHPICQHGMEADADNEVLIW  
RFDSSLARRHRIARELHPDFYKDC\$

Fig. 106B

## 2003 CON 14 BG nef. OPT

ATGGCGGCAAGTGGTCCAAGTGTCCATCGTGGGCTGGCCCGAGGTGGCGGAGCGCATCCGCCGACACCCCGCCCGCGCGGTGGG  
CGCCGTGTCCAGGACCTGGCCAAGCACGGCGCATCACTCTCCAACACCGCGCCCAACAACCCGACTGCGCTGGCTGGAGGCCAGG  
AGGAGGACTCCGAGGTGGGCTTCCCGTGGCCCTCCAGGTGCCCCATGACCTACAGGGCGCTTCGACCTGTCTTCTCCTG  
AAGGAGAAGGGCGGCTGGACGGCTGATCTACTCAAGCAGCGCCAGGACATCTGGACCTGTGGGTGTACAACACCCAGGCTTCTCCC  
CGACTGGCAGAACTACACCCCGGCGCCGCTACCCCTGACCTTCGGCTGGTGTCTCAAGCTGGAGCCCGTGGACCCCGCGGAGG  
TGAGGAGGCCACCAAGGGCGAGAACATCCCTGTGCAACCCCATCTGCCAGCACGGCATGGAGACGCCGACACGAGGTGCTGATCTGG  
CGCTTCGACTCCTCCTGGCCCGCGCCACATCGCCCGGAGCTGCACCCCGACTTCTACAAGGACTGCTAA

Fig. 107A

## 61. 2003 2003 CON s pol. PEP

FFRENLAFOQGEAREFSSEQTRANSPTSRELVRGGDNPLSEAGAEQGTVSLSFPOITLWQRPPLVTVKIGGQLKEALLDTGADDTVLEEIN  
LPKWKPKWIGGIGGFIVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPKPGMDGPKVKQWPLTEEK  
IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKDDSTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDE  
DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGPSPAI FQSSMTKILEPFRQNPFIYQYMDLLYVGSDEIGQHRTKIEELREHLLRWGF  
TTPDKKHQKEPPFLWMGYELHPDKWTVPQIQLPEKDSWTVNDIQKLVGKLNWASQIYPGKVKQLCKLLRGAKALTDIVPLTEEALELEAEN  
REILKEPVHGVYDPSKDLIAEIQKQGDQWTYQIYQEPFKNLKTGYAKMRSHTNDVKQLTEAVQKIATESIVWGKTPKFRPLPIQKETW  
ETWWTEYWQATWIPWEFFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRQKQVSVLTETTNQKTELQAIHLALQDSG  
SEVNIIVTDSQYALGIIQAPDKSESELVNIIEQLIKKEKVLWSVPAAHKGIGGNEQVDKLVSTGIRKVLFDGIDKAQEEHEKYHSNWRAM  
ASDFNLPPIVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEKIIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH  
TDNGSNFTSAAVKAACWAGIQQEFGIPYNPQSGVVESMNKELKKIIGQVRDQAEHLKTAQMAVFIHNFRRKGGIGGYSAGERIIDIIAT  
DIQTKELQKQITKIQNFVRYRDSRDPINWGPAPKLLWKGEAVVIQDNSEIKVVPRRKAKIIRDYGKQMAAGDDCVAGRQDEDS\$

## Fig. 107B

2003 CON S pol. OPT

TTCTTCGGGAGAACCTGGCCCTCCAGCAGGGGAGGCCCGGAGTTCTCTCCGAGCAGACCCGGCCAACTCCCCAACCTCCCGGAGCTGCGCGTGCGG  
 CGGGGGGACAAACCCCTGTCCGAGGGCCGGCCGAGCGCCAGGACCCGTGTCCCTGTCTTCCCCAGATCACCTTGTGGCAGCGCCCTTGGTACCCG  
 TGAAGATCGGGCCAGCTGAAGAGGCCCTGTGTGACACCGGCGGACGACACCGTGTGGAGGAGATCAACCTGCCCCGGCAAGTGAAGCCCAAGATG  
 ATCGGGCGCATCGGGCGCTTCATCAAGGTGCGCCAGTACGACCAAGATCTGTATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGTGTGGGCCCCAC  
 CCCCCTGAACATCATCGGGCCGAACATGCTGACCCAGATCGGCTGACCCCTGAACCTTCCCCATCTCCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCG  
 GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCAECGAGATGGAGAAGGAGGCAAGATCTCC  
 AAGATCGGGCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAAGAGAAGGACTCCACCAAGTGGCGAAGCTGGTGGACTTCCCGGAGCTGAACAA  
 GCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGGGCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGGCGACGCCCTACT  
 TCTCCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCCCTTACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACAGTACAACGTGCTG  
 CCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCTTCCATGACCAAGATCTTGAGGCCCTTCCGCACCCAGAACCCCGAGATCGTGATCTACCAGTA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGGAGCACCTGCTGCGCTGGGGCTTCAACACCC  
 CCGACAAGAAGCACCAAGAGGCCCTTCTGTGGATGGGCTAGAGCTGACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGAGAAGGAC  
 TCCTGGACCGTGAACGACATCCAGAACGTGGTGGCAAGCTGAACCTGGGCCCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGTCTGG  
 CGGGCCAAAGGCCCTGACCGACATCGTGGCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCCTGCACGGCGTGT  
 ACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGACCAAGTGGACCTTCCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGATGGCTCCGCCACACCAACGACGCTGAAGCAGCTGACCGAGGCCGTGCAAGAGATCGCCACCGAGTCCATCGTATCTGGGGCAA  
 GACCCCAAGTTCCGCCCTGCCATCCAGAAGGAGACCTGGGAGACCTGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCTTGTGTAAGTGTGGTACCACTGGAGAACGAGCCCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCCCGCCAAACCGCGAGACCAAGCTG  
 GGCAAGGCCGGCTAGTGACCGACCGCGGGCCGCAAGAGTGGTTCCTTGACCGGATCCAGTACGCCCTGGGCATCATCCAGGCCCGAGCCGACAAGTCCGAGTCCGAGCTGGTGAACC  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCCGAGCCGACAAGTCCGAGTCCGAGCTGGTGAACC  
 AGATCATCGAGCAGTGTATCAAGAAGGAGAAGGTGTACCTGTCTTGGTGGCCGCCACCAAGGGCATCGCGGGCAACGAGCAGGTGGACAAGCTGGTGTCC  
 ACCGGCATCCGCAAGTGTCTTCCCTGGACGGCATCGACAAGGCCGAGGAGGACGAGAGTACCACTCCAACCTGGCGGCCATGGCCCTCCGACTTCAA  
 CCTGCCCCCATCGTGGCCAAAGGATCGTGGCCCTTCTGTGGCCGTGCACGTGGCCCTCCGGCTACATCGAGGCCGAGGTGATCCCGCGCGAGACCCGGC  
 GGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGGCCGTGCACGTGGCCCTCCGGCTACATCGAGGCCGAGGTGATCCCGCGCGAGACCCGGC  
 CAGGAGACCGCCCTACTTATCTTGAAGCTGGCCCGCCCTGGCCCTGAAGGTGATCCACACCGACAACCGGCTCCAACCTTACCTCCGCCCGCCGTGAAGGC  
 CGCCTGTGTGGCCGGCATCCAGCAGGAGTTCCGCATCCCTACAACCCAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTTCGGACCCAGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGGCATCGGGCGGTACTCC  
 GCGGGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGTGCAGAAGCAGATCACCAAGATCCAGAACCTTCCGGCTGTACTACCGCGA  
 CTCCCGGACCCCATCTGGAAGGGCCCGCCAAAGCTGTGTGGAAGGGCGAGGGCGCGGTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGGTCCCCGGC  
 GCAAGGCCAAGATCATCCGGGACTACGGCAAGGAGATGGCCGGCGCCCGCCAGGACGAGGACTAA

141/178

Fig. 108A

## 62 2003 M GROUP anc pol. PEP

FFRENLAQQGEAREFSSEQTRANSPTSRDLWDGGRDLSLPEAGAERQGTVSFSPQITLWQRPVLTIKIGGQREALLDTGADDTVLEEIN  
 LPGAQKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLGPPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKDKSTKWRKLVDFRELNKRQTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDE  
 DFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFTKNPEIYIYQYMDLTVGSDLEIGQHRAKIEELREHLLRWGF  
 TTPDKKHQKEPFLWMGYELHPDKWTVPQIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAEELELAEN  
 REILKEPVHGVYDPSKDLIAEIQKGQDQWTYQIYQEPFNKLTGKYAKMRSATNDVKQLTEAVQKIATESIIVWGKTPKFRLPQKETW  
 ETWTEYWOATWIPWEFEVNTPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRQKVVSLETETTNQKTELQAIHLALQDSG  
 SEVNIIVTDSQYALGIIQAQPKSESELVNIQIEQLIKKEKYLWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAM  
 ASDFNLPVVAKEIVASCDCQKLGAMHGQVDCSPGIWQLDCTHLEKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH  
 TDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVMQMAVFIHNFKRKGGIGGYSAGERIIDIIAT  
 DIQTKELQKQITKIQFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSEIKVVPRRKAKIIRDYGKQMGAGDDCVAQRQDEDS

Fig. 109A

## 63. 2003 CON A1 pol. PEP

FFRENLAQQGEAREFSSEQTRANSPTSRDLWDGGRDLSLPEAGAERQGTGPTFSFPQITLWQRPVLTIVRIGGQKEALLDTGADDTVLEDI  
 NLPQKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLGPPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEE  
 KIKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKDKSTKWRKLVDFRELNKRQTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLD  
 ESFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRSKNPEIYIYQYMDLTVGSDLEIGQHRTKIEELRAHLLSWG  
 FTTPDKKHQKEPFLWMGYELHPDKWTVPQIQLPEKESWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEAEELELA  
 NREILKDPVHGVYDPSKDLIAEIQKGQDQWTYQIYQEPFNKLTGKYARKRSATNDVKQLAEVVKVVMESIVIWGKTPKFKLPQKET  
 WETWMDYWOATWIPWEFEVNTPLVKLWYQLEKDPVGAETFYVDGAANRETKLGKAGYVTDGRQKVVSLETETTNQKTELQAIHLALQDS  
 GSEVNIIVTDSQYALGIIQAQPKSESELVNIQIEKLIGKDKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWR  
 MASDFNLPPIVAKEIVASCDCQKLGAMHGQVDCSPGIWQLDCTHLEKVLVAVHVASGYIEAEVIPAETGQETAYFLIKLAGRWPVKV  
 HTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVMQMAVFIHNFKRKGGIGGYSAGERIIDIIA  
 TDIQTKELQKQITKIQFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSEIKVVPRRKAKIIRDYGKQMGAGDDCVAQRQDEDS

142/178

## Fig. 108B

2003\_M\_GROUP and pol. OPT

TTCTTCGGGAGAACTGGCCTTCCAGCAGGGCGAGGGCCCGGAGTTCTCCTCCGAGCAGACCCGGCCCAACTCCCCACCTCCCGCGAGCTGCGCGTGG  
 CGCGCGGACAACCCCTGTCCGAGGCGGGCGGAGCGCCAGGACCGTGTCTTCTCTTCCCCCAGATCACCTGTGGCAGGCCCCCTGGTGACCA  
 TCAAGATCGGCGGCAGCTGCGGAGGCCCCGTGTGGACACCGGCGCGACACCGTGTGGAGGAGATCAACCTGCCGCAAGTGAAGCCCAAGATG  
 ATCGGCGGCATCGGCGGCTTCATCAAGTGGCGCCAGTACGACCAAGATCCTGATCGAGATCTGGGCAAGAGGCCATCGGCACCGTGTGGTGGGCCCCAC  
 CCCCCTGAACATATCGGCGGCAACATGCTGACCCAGATCGGCTGACCCCTGAACTTCCCCATCTCCCCATCGAGACCGTGGCCGTGAAGCCCCG  
 GCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGAGATCTGACCCGAGATGGAGAGGCAAGATCTCC  
 AAGATCGGCCCCGAGAACCCCTACACACCCCGTGTTCGCCATCAAGAAGAAGACTCCACCAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAA  
 GCGACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCACCCCGCGCTGAAGAAGAAGAAGTCCGTGACCGTGGACGTGGCGGACGCGCTACT  
 TCTCCGTGCCCTTGACGAGGACTTCCGCAAGTACACCGCCTTACCATCCCCCTCCATCAACAAACGAGACCCCGGCATCCGCTACAGTACAACGTGCTG  
 CCCCAGGGCTGGAAGGGCTCCCCCGCATCTTCCAGTCTCCATGACCAAGATCCTGGAGCCCTTCCGCACCAAGAACCCCGAGATCGTGTATCTACCAAGTA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGACCCGCCAAGATCGAGGAGCTGGCGGAGCACTGTGCGTGGGGCTTCAACACCC  
 CCGACAAGAAGCAGACGAGGACCCCTTCTGTGGATGGGTACGAGTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGAGAAGGAC  
 TCCTGGACCGTGAACGACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCTTCCAGATCAACCCCGCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGG  
 CGCGCCCAAGGCCCTGACCGACATCGTGCCCTGACCGAGGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCCGTGCACGCGGTGT  
 ACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGACCACTGGACCTACCAAGATCTACAGGAGCCCTTCAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGATGCGCTCCGCCACACCAACGACGTGAAGCAGTGAACGAGGCGGTGCAGAAGATCGCCACCGAGTCCATCGTATCTGGGGCAA  
 GACCCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCCGTGAACA  
 CCCCCCTTGGTGAAGCTGTGTACCAAGCTGGAGAAGGAGCCCATCGTGGGCGCGGAGACCTTCTACGTGGACGGCGCGCCCAACCGCGAGACCAAGCTG  
 GGCAAGGCGCGCTACGTGACCGACCGGCGCGCCAGAAAGTGGTGTCCCTGACCGAGACCCACCAACCAAGAACCCGAGCTGCAGGCCATCCACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCACGCCGACAAGTCCGAGTCCGAGCTGGTGAACC  
 AGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGTCTGGTGGCCGCCACAAAGGCGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTGTCC  
 TCCGGCATCCGCAAGGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGGAGCAGAGAAGTACCACTCCAACCTGGCGGCCCATGGCCTCCGACTTCAA  
 CCTGCCCCCGTGGTGGCCAAAGGATCGTGGCCCTCCTGCGACAAGTGCAGCTGAAGGCGGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCT  
 GGCAGCTGGAATGCAACCACTGGAGGGCAAGGTGATCTGTGGTGGCGGTGCAGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACCGGC  
 CAGGAGACCGCTACTTCATCTGAAGCTGGCCGCGCGCTGGCCCGGTGAAGGTGATCCACACCGACAACCGGTCCAACCTCACCTCCGCCGCGGTGAAGGC  
 CGCTGTGTGGCGCGGCATCCAGCAGGAGTTCGGCATCCCTACAAACCCCAAGTCCAGGCGGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCGACCAAGCCGAGCACCTGAAGACCGCGCTGCAGATGGCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGGCATCGGCGGCTACTCC  
 GCCGCGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCCAGAAGTTCGCGGTGTACTACCGCGA  
 CTCGCGGACCCCATCTGGAAGGGCCCCCGCAAGCTGTGTGGAAGGGCGAGGGCGCGCTGGTGTATCCAGGACAACTCCGAGATCAAGGTGGTGGCCCCGC  
 GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCCGCGCGGACGACTGCGTGGCGCGCGCGCGGACGAGGACTAA

143/178







145/178

## Fig. 109C

64. 2003 A1.anc pol.PEP

FFRENLAFOQGEARKFSSEQTRANSPTSRELWDGGRDLSLLSEAGAERQGTVPFSFPQITLWQRP LVTVKIGGQLKEALLDTGADDTVLEDI  
NLP GKWKPKMIGGIGGFVKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEE  
KIKALTEICTEMEKEGKISKIGPENPNYNTPVFAIKKKDSTKWRKLVDFRELNRKTQDFEWVQLGIPHPAGLKKKSVTVLVDGDYFVSPLD  
ESFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRSKNPEIVIYQYMDDLVVGSDLEIGQHRAKIEELRAHLLSWG  
FTTPDKKHQKEPPFLWMGYELHPDKWTVPQIKLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEALELAE  
NREILKDPVHGYYDPSKDLVAEIQKGQDQWTYQIYQEPFNKLTGKYAKKRSHTNDVKQLTEVVQKVATESIVIWGKTPKFRLP IQKET  
WETWMEYWQATWIPWEFEVNTPLVLWYQLEKEPIAGAETFYVDGAANRETCLGKAGYVTDGRQKVVSLETETTNQKTELHAIHLALQDS  
GSEVNI VTD SQYALGI IQAOPDRSESELVNQIIEKLLIEKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
MASDFNLPPIVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKV  
HTDNGSNFTSAAVKAACWWANIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNEFRKGGIGGYSAGERIIDIIA  
TDIQTKELOKQITKIQNFRVYRDSRDPINWKGPAKLLWKGEAVVIQDNSDIKVVP RRKAKIIRDYGMAGDDDCVAGRQDED\$

146/178

## Fig. 109D

2003\_A1.anc pol.OPT

TTCTTCCGGAGAACCTGGCCCTCCAGAGGGCGAGGCCCGCAAGTTCTCTCCGAGCAGACCCGCGCAACTCCCCACCTCCCGGAGCTGTGGGACGG  
 CGGCCGGACTCCCTGCTGTCGAGGCGCGCGCGAGCGCCAGGCAACGTCCTCTCTCTCCCCAGATCACCTGTGGCAGCGCCCTGGTGA  
 CCGTGAAGATCGGCGGCGAGCTGAAGAGGCCCTGCTGGACACCGGCGCGACGACACCGTCTGGAGGACATCAACCTGCCGCAAGTGAAGCCCAAG  
 ATGATCGGCGGATCGGGGCTTCATCAAGGTGGCCAGTACGACACAGATCCTGATCGAGATCTGGGCAAGAGGCCATCGGCACCTGCTGGTGGGCC  
 CACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGATCGGCTGACCCCTGAACCTTCCCATCTCCCCATCGAGACCGTGCCTGTAAGCTGAAGC  
 CCGCATGGACGSCCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGACCCGAGATGGAGAGGGCAAGATC  
 TCCAAGATCGGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAA  
 CAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCACCCCGGCTGAAGAAGAAGTCCGTGACCGTCTGGACGTGGCGGACCGCT  
 ACTTCTCCGTGCCCTGGACGAGTCTTCCGCAAGTACACCGCTTCAACATCCCTCCATCAACAACGAGACCCCGCATCCGCTACCAACAGTG  
 CTGCCACAGGCTGGAAGGCTCCCCCGCATCTTCCAGTCTCCATGACCAAGATCCCTGGAGCCCTTCGGTCCAAGAACCCCGAGATCGTGATCTACCA  
 GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCGCAAGATCGAGGAGCTGCGCGCCACCTGTCTCTGGGGCTTCACCA  
 CCCCAGACAAGAACCAAGAGGCCCTTCTGTGGATGGCTACGAGCTGCACCCCGACAAAGTGGACCGTGCAGCCCATCAAGCTGCCCGAGAAG  
 GACTCTGGACCGTGAACGACATCCAGAACCTGGTGGCAAGCTGAAGTGGCTCCAGATCTACCGCGGATCAAGTGAAGCAGCTGTGCAAGCTGCT  
 GCGGGCGCAAGGCCCTGACCGACATCTGACCTGACCGAGGCGGAGCTGGAGCTGGCCGAGAACCCCGGAGATCTTGAAGGACCCCTGCGACGGG  
 TGTACTAGACCCCTCCAAGACCTGGTGGCCGAGATCCAGAACGAGGCCAGGACCAAGTGGACCTACCGAGATCTACCGAGGCCCTTCAAGAACCTGAAG  
 ACCGCAAGTACGCCAAGAGCGCTCCGCCACACCAACGACGTGAAGCAGTGAAGCAGTGAAGTGGTGGTGGATCTGGCAGGCCACTGGATCCCCGAGTGGAGTCTGTGA  
 CAAGACCCCAAGTTCGGCTGCCATCCAGAGGAGACCTGGGACCTGGTGGATGGAGTACTGGCAGGCCACTTCTACGTGGACGGCGCCCAACCGCGAGACCAAG  
 ACACCCCGCTGGTGAAGCTGTGGTACCACTGTGACCGACTCCAGTACCGCTGGCATCATCAGGCCAGCCGACCAACCGAGACCAACCGAGACCGAGCTGCACGCCATCCACCTGGC  
 CTGGCAAGGCGGCTACGTGACCGACCGGCGCGCCAGAGGTGGTGTCCCTGACCGAGACCAACCGAGACCGAGCTGCACGCCATCCACCTGGC  
 CCTGCAGGACTCCGGCTCCGAGTGAACATCGTGACCGACTCCAGTACCGCTGGCATCATCAGGCCAGCCGACCGCTCCGAGTCCGAGCTGGTGA  
 ACCAGATCATCGAGAAGCTGATCGAGAAGGAGAGGTGTACCTGTCTGGTGGCGCCACAAAGGCAACGAGGATCGGGCGCAACGAGCAGGTGGACAAGCTGGTG  
 TCCTCCGGCATCCGCAAGGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGGACGAGAAAGTACACTCCAACCTGGCGGCCCATGGCTCCGACTT  
 CAACCTGCCCCCATCGTGGCCAAAGGATCGTGGCTCTCTGGACAAGTGGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
 TCTGGCAGTGGACTGCACCCACCTGGAGGCAAGGTGATCTGTGGTGGCTGCAGTGGCTCCGGTACATCGAGGCCGAGGTGATCCCCGCGGAGACC  
 GGCCAGGAGACCGCTACTTCTGTGAAGTGGCCGCGCTGGCCGCTGAAGTGGTGCACACCGACAAAGGCTCCAACCTTCACTCCGCGCGCGTGA  
 GGCGGCTGCTGGTGGCCCAACATCCAGCAGGATTCGGCATCCCTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
 TCATCGGCCAGGTGCGGAGCAGGCGGACCTGAAGACCGCGTGCAGATGGCCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGCATCGGCGGCTAC  
 TCCGCGGCGGAGGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGTGCAGAAGCAGATCACCAAGATCCAGAACCTTCCGCGTGTACTACCG  
 CGACTCCCGGACCCCATCTGGAAGGGCCCCGCAAGTGTGTGGAAGGGCGAGGGCGCGTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGCCCC  
 GCCGCAAGGCCAAGATCATCCGCAAGTACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGCGCCAGGACGAGGACTAA

Fig. 110A

## 65. 2003 CON A2 pol. PEP

FFRENLAQQREARKFSSEQNRANSPTSRELNRNGGRDNLSEAGAEQGVHSCNFPQITLWQRPLVTVKIEGQLREALLDTGADDDTVLEDI  
 NLPGRWPKMIGGIGGFIKVRQYDQIAIEICGKRAIGTVLVGTPVNI IGRNMLVQLGCTLNFPI SPIETVPVKLPGMDGPKVKQWPLTEE  
 KIKALTEICKEMEKEGKISKIGPENPYNTPVFAIKKDDSTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTVLVDVGDAYFVSPLH  
 EDFRKYTAFTIP SINNETPGIRYQYNVLPQGKGSPIFQSSMTKILEPFRSKNPEMVIYQYMDLTVVGSDEIGQHRAKIEELRAHLLRWG  
 FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTDIVTLTKEAELELEE  
 NREILKNPVHGVYDPSKDLIAEIQKQGQDQWTYQIYQEPFNKLTGKYAKRKSTHTNDVKQLTEAVQKIAIESIVIWGKTPKFRLP IQKET  
 WETWTEYWQATWIP EWEFVNTPLVLKLYQLETEPIAGAEFFYVDGAANRETKLGKAGYVTDGRQKIVSLTETTNOKTELHAIYLAQDS  
 GLEVNIVTDSQYALGIIQAQPDKSESELVNIIEKLIIEKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEHYHSNWRA  
 MAHDFNLPPIVAKEIVASCDKQKLGEMHGQVDCSPGIWQDCTHLEKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI  
 HTDNGPNFTSATVKAACWAGVQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA  
 TDIQTKELQKQIITKIQNFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAIIIRDYGMAGDDCVCAGRQDED\$

Fig. 111A

## 66. 2003 CON B pol. PEP

FFREDLAFQCKAREFSSEQTRANSPTRRELQVWGRDNNLSLSEAGADRQGVTSFSPQITLWQRPLVTIKIGGQLKEALLDTGADDDTVLEEM  
 NLPGRWPKMIGGIGGFIKVRQYDQIILIEICGKHAIGTVLVGTPVNI IGRNLLTQIGCTLNFPI SPIETVPVKLPGMDGPKVKQWPLTEE  
 KIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDDSTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTVLVDVGDAYFVSPLD  
 KDFRKYTAFTIP SINNETPGIRYQYNVLPQGKGSPIFQSSMTKILEPFRKQNPDIYIYQYMDLTVVGSDEIGQHRTKIEELRQHLLRWG  
 FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVIPLTEAELELAE  
 NREILKEPVHGVYDPSKDLIAEIQKQGQDQWTYQIYQEPFNKLTGKYARMGAHTNDVKQLTEAVQKIAIESIVIWGKTPKFRLP IQKET  
 WEAWTEYWQATWIP EWEFVNTPLVLKLYQLEKEPIVGAETFFYVDGAANRETKLGKAGYVTDGRQKVVSLTDTTNOKTELQAIHLALQDS  
 GLEVNIVTDSQYALGIIQAQPDKSESELVSIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGDIDKAQEEHEHYHSNWRA  
 MASDFNLPPVVAKEIVASCDKQKLGEMHGQVDCSPGIWQDCTHLEKIIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTI  
 HTDNGSNFTSTTVKAACWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIA  
 TDIQTKELQKQITKIQNFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAIIIRDYGMAGDDCVCAGRQDED\$

147/178

Fig. 110B

2003\_CON\_A2\_pol.OPT

TTCTTCGGGAGAACCTGGCCCTCCAGCAGCGCGAGGGCCCGCAAGTTCTCTCCGAGCAGAACCGGGCCCAACTCCCCACCTCCCGGAGCTGGCGAACGG  
 CGCCCGGACAACTGCTGTCGAGGGCCGGCCGAGGAGCAGGGCACCGTGACTCTCTGCAACTTCCCCAGATCACCTGTGGCAGCGCCCTTGGTGA  
 CCGTGAAGATCGAGGGCCAGCTGCGCGAGGGCCCTGCTGGACACCGGCGCGACGACACCGTGTGGAGGACATCAACCTGCCGGAAGTGAAGCCCAAG  
 ATGATCGGGGCGATCGGGCGCTTCATCAAGGTGCGCCAGTACGACAGATCGCCATCGAGATCTGCGGCAAGCGGCCATCGGCACCGTCTGTTGGGCC  
 CACCCCGTGAACATCATCGGGCGCAACATGCTGTGAGCTGGCTGACCTGAACTTCCCCATCTCCCCATCGAGACCGTGGCCGTGAAGCTGAAGC  
 CCGCATGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGAAGGAGATGGAGAGGAGGCAAGATC  
 TCCAAGATCGGGCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAGCTGAA  
 CAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCGGCTGAAGAAGAAAGTCCGTGACCGTGTGGACGTGGGCGGACGCGCT  
 ACTTCTCCGTGCCCTGCACGAGGACTTCCGCAAGTACACCGCCCTTACCATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACCATCAACACGCTG  
 CTGCCCCAGGGCTGGAGGGCTCCCCCGCATCTTCCAGTCTCCATGACCAAGATCTCTGGAGCCCTTCCGCTCCAAGAACCCCGAGATGGTGATCTACCA  
 GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGCTGCGGCGCCACCTGCTGCGTGGGGCTTCACCA  
 CCCCCAAGAAGCACAGAGAGGAGCCCCCTTCTGTGGATGGCTACGAGTGGACCCCGACAAAGTGGACCGTGCAGCCATCAAGCTGCCCGAGAAG  
 GACTCTGACCCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAAGTGGGCTCCAGATCTACCGCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCT  
 GCGGGCACCAAGGCCCTGACCGACATCGTGACCTTGACCAAGGAGGCCGAGCTGAGCTGGAGGAGAACCGCGAGATCTTGAAGAACCCCGTGCACGCGG  
 TGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGACCAAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAG  
 ACCGGCAAGTACGCCAAGTCCACCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAAGATCGCCATCGAGTCCATCGTGATCTGGGG  
 CAAGACCCCAAGTTCGGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCTGTGA  
 ACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGACCGAGCCCATCGCCGGCGGCGAGACCTTCTACGTGGACGGCGCCCAACCGCGAGACCAAG  
 CTGGCAAGGGCCGGCTACGTGACCGACCGCGCGCCGCGAGAGATCGTGTCTTACCGAGACCCACCAACAGAGACCGAGCTGCACGCCATCTACCTGGC  
 CCTGCAGGACTCCGGCTTGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCCGACGCCGACCGCTCCGAGTCCGAGCTGGTGA  
 ACCAGATCATCGAGAAAGTATCGAGAAAGGAGCGCGTGTACCTGTCTTGGGTGCCCGCCACAAGGGCATCGCGGGCAACGAGCAGGTGGACAAAGCTGGTG  
 TCCCTCCGGCATCCGCAAGTGTCTTCCCTGGACGGCATCGACAAGGCCCGAGGAGCACGAGCGCTACCACTCCAACCTGGCGGCCATGGCCACGACTT  
 CAACCTGCCCCCATCGTGGCCAAAGGAGATCGTGGCCCTCTGCGACAAGTCCAGCTGAAGSGGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
 TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGGTGATCTGTGTGGCGTGCACGTGGCTTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACC  
 GGCCAGGAGACCGCCCTACTTCATCTGAAGCTGGCCGGCCGCTGGCCCGTGAAGGTGATCCACACCGACAAACGGCCCAACTTCACTCCGCCACCGTGAA  
 GGCGGCTGTGTGGCGGCGGTGCAGCAGGAGTTCGGCATCCCCACAACCCCGATCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
 TCATCGGCCAGGTGCGGACCGAGCCGACCTGAAGACCGCGCTGCAGATGGCCGTTCATCCACAACCTCAAGCCCAAGGGCGCATCGGCGGCTAC  
 TCCGCGGCGAGCGCATCATCGACATCATCGCACACCGACATCCAGACCAAGGAGCTGCAGAAAGCAGATCATCAAGATCCAGAACTTCCGCGTGTACTACCG  
 CGACTCCCGGACCCCATCTGAAGGGCCCGCAAGCTGTGTGGAGGGCGAGGGCGCGTGGTGTATCCAGGACAACCTCCGACATCAAGGTGGTGGCCCC  
 CCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGCGACGACTGCGTGGCCGGCGCGCAGGACCGAGGACTAA

148/178

## Fig. 111B

2003\_CON\_B\_pol.OPT

TTCTTCCGCGAGGACCTGGCCCTTCCCCAGGGCAAGGCCCGGAGTTCTCTCCGAGCAGACCCGGCCAACTCCCCCACCCTCCGCGAGCTGCAGGTGTG  
 GGGCCGCGACAACAACCTCCCTGTCCGAGGCGGCGGACCGCCAGGGCACCGTGTCTCTTCCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGA  
 CCATCAAGATCGGCGGCCAGCTGAAGAGGCCCTGCTGGACACCGGCGCCGACACCGTGTGGAGGATGAACCTGCCCCGCCCTGGAAGCCCCAAG  
 ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCGCTACGACCAAGATCTCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGTGGTGGCCC  
 CACCCCGTGAACATCATCGGCGCAACCTGCTGACCCAGATCGGCTGCACCTGAACCTCCCCATCTCCCCCATCGAGACCGTGCCTGGAAGCTGAAGC  
 CCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC  
 TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAGGACTCCACCAGTGGCGCAAGCTGGTGGACTTCGCGGAGCTGAA  
 CAAGCGACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCCACCCTCCGCTGAAGAAGAAAGTCCGTGACCGTGTGGACGTGGCGGACGCT  
 ACTTCTCGTGGCCCTGGACAAGGACTTCGCAAGTACACCGCTTCACCATCCCTCAACACGAGACCCCCCGCATCCGCTACCACTACCAACAGCTG  
 CTGCCCCAGGGCTGGAAGGCTCCCCCGCATCTTCCAGTCTCCATGACCAAGATCTGGAGCCCTTCGCGCAAGCAGAACCCCGACATCGTGTATCTACCA  
 GTACATGGACGACCTGTACGTGGGCTCCGACTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGTGGCCAGCACCTGCTGCGTGGGGCTTCACCA  
 CCCCCACAAGAAGCACCAAGAGGAGCCCCCTTCTGTGGATGGGTACGAGCTGCACCCGACAAAGTGGACCGTGCAGCCCCATCGTGTGCCCGAGAAG  
 GACTCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAAGTGGCTCCAGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCT  
 GCGGGACCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCCCGAGATCTCTGAAGGAGCCCCGTGCACGGCG  
 TGTACTACGACCCCTTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGCTGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAG  
 ACCGCAAGTACGCCCGCATCGCGGCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATGCCACCGAGTCCATCGTGTATCTGGGG  
 CAAGACCCCAAGTTCAAGCTGCCCCATCCAGAAGGAGACCTGGAGGCCCTGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGA  
 AGACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAAGGAGCCCATCGTGGGCGCGGAGACCTTCTACGTGGACGGCGCCGCAACCCGAGACCAAG  
 CTGGGCAAGGCCGGCTACGTGACCGACCGCGCGCGCCGCAAGAGTGGTGTCCCTGACCGACACCAACCGAGACCGAGCTGCAGGCCATCCACCTGGC  
 CCTGCAGGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAAGTCCGAGTCCGAGCTGGTGT  
 CCCAGATCATCGAGCAGCTGATCAAGAAGAGAAGTGTACCTGGCTGGGTGCCCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTG  
 TCCGCCGGCATCCGCAAGGTGTCTTGACCGCATCGACAAGGCCCAAGGAGGACGAGAAATACCACTCCAACTGGCGCGCCATGGCTCCGACTT  
 CAACCTGCCCGCTGGTGCCAAAGGAGATCGTGGCTCTCTGCGACAAGTGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
 TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGGTGGCTGCACGTGGCTCCGGTACATCGAGGCCGAGGTGATCCCCGCGAGACC  
 GGCCAGGAGACCGCTACTTCTGTGAAGCTGGCGCGCTGGCCCGTGAAGACCATCCACACCGACAACGGCTCCAATTCACTCCACCAACCGTGAA  
 GGCGGCTGTGTGGCGCGCATCAAGCAGGAGTTCGGCATCCCCTACAACCCAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGA  
 TCATCGGCCAGGTGGCGGACCGAGCACCCTGAAGACCGCGCTGCAGATGGCCGTGTTCATCCACAATTCAGCGCAAGGGCGGCATCGGCGGCTAC  
 TCCGCCGGGAGGCGCATCGTGGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCAACCAAGATCCAGAACTTCCGCGTGTACTACCG  
 CGACTCCCGGACCCCTGTGGAAGGGCCCCGCAAGCTGTGTGGAAGGGCGAGGGCGCGTGGTGTATCCAGGACAACTCCGACATCAAGGTGGTGGCCCC  
 GCGGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGCGACGACTGCGTGGCTTCCCGCAGGACGAGGACTAA

149/178

150/178

## Fig. 111C

67. 2003 B.anc pol. PEP

FFRENLAFFQCKAREFSSEQTRANSPTRRELQVWGRDNNPLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEM  
NLPCKWKPKMIGGIGGFIKVRQYDQILLIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEE  
KIKALVEICTEMEKEGKISKIGPENPNYTPVFAIKKKDSTKWRKLVDFEELNKRRTQDFWEVQLGIPHAPAGLKKKKSVTVLDVGDAYFSPVPLD  
KDFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPFIYIYQYMDLTVGSDLEIGQHRKIEELREHLLRWG  
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPLTEEALELAE  
NREILKEPVHGVYDPSKDLIAEIQKQGQGWYQIYQEPFKNLKTGKYARMRGHTNDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKET  
WEAWTEYWQATWIPWEFVNTPLVLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRQKVVSITDTTNQKTELQAIHLALQDS  
GLEVNIVTDSQYALGIIQAOPDKSESELVSQIIQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
MASDFNLPPVVAKEIVASCDCQKLGAEAMHGQVDCSPGIWQLDCTHLEKIIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI  
HTDNGSNFTSTTVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIA  
TDIQTKELOKQITKIQNFRVYRDSRDLWKGPALKLWKGEAVVIQDNSDIKVVPRRKAKIIRDYKGQMGADDCVASRQDED\$



152/178

## Fig. 112A

## 68. 2003 CON C pol.PEP

FFRENLAFFQGEAREFPSEQTRANSPSTRSRELQVRGDNPRSEAGAERQGTNLNFPQITLWQRPLVSIKVGQIKEALLDTGADDTVLEEINLPG  
KWKPKMIGGIGGFIVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKA  
LTAICEEMEKEGKITKIGPENPYNTPVFAIKKSDSTKWRKLVDFRELNKRTOQDFWEVQLGIPHPAGLKKKSVTVLVDGDYFVSVPLEDEGFR  
KYTAFTIPSINNETPGIRYQYNVLPQGWKSPAFQSSMTKILEPFRAQNPEIYIYQYMDLLYVGSDEIGQHRAKIEELREHLLKKGFTTP  
DKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQLVGKLNWASQIYPGKVRQLCKLLRGAKALTDIVPLTEEALELAENREI  
LKEPVHGVYDPSKDLIAEIQKQGHQDWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIWIWGTTPKFRLPQKETWETW  
WTDYWOATWIPWEFVNTPLVLKLYOLEKEPIAGAETFYVDGAANRETAKIGKAGYVTDGRQKIVSLTETTNQKTELQAIQALQDSDGSEV  
NIVTDSQYALGIIQAQPKSESELVNIIEQLIKKERVYLSWPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASE  
FNLPPIVAKEIVASCDKCOLKGEAIIHGQVDCSPGIWQDCTHLEGKIIIVAVHVASGYIEAEVIPAETGOETAAYIILKLAGRWPVKVIHTDN  
GSNFTSAVKAACWWAGIQOEFGIPYNPQSGVVESMNKELKKIIGQVRDOAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDLIATDIQ  
TKELQKQIIKIQNFRVYRDSRDP IWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIKDYGKQMGACDCVAGRQDED\$



Fig. 112B

2003 CON C pol.OPT

[illegible]

154/178

## Fig. 112C

## 69. 2003 C.anc pol.PEP

FFRENLAFFQGEAREFPSEQTRANSPTSRELQVGRDNPRSEAGAERQGTLLNFPQITLWQRPLVSIKVGQIKEALLDTGADDTVLEEINL  
 PGKWKPKMIGGIGGFIKVRQYDQILLIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKI  
 KALTAICEEMEKEGKITKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKRQTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDEG  
 FRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFRAQNPEIYIYQYMDLLYVGSDLEIGQHRAKIEELREHLLKWGFT  
 TPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEALELELAENR  
 EILKEPVHGVYDPSKDLIAEIQKQGHQWTVQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIWIWGTTPKFRLP IQKETWE  
 TWWTDYWQATWIPWEFVNTPLVLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTDGRQKIIVSLTETTNQKTELQAIQALQDSGS  
 EVNIVTDSQYALGIIQAQPKSESELVNQIIQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAQEEHEKYHSNWRAMA  
 SEFNLPPIVAKEIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEKIIIVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHT  
 DNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVMQMAVFIHNFKRKGIGGYSAGERIIDIIATD  
 IQTKELQKQIIKIQNFRVYYRDSRDP IWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIRDYGKQMGADCVAGRQDED\$

## Fig. 112D

2003\_c\_anc pol.1.OPT

TTCTTCCGCGAGAACCTGGCCCTTCCCCAGGGCGAGGCCCGGAGTTCCCTCCGAGCAGACCCGCGGCAACTCCCCCACTCCCCGCGAGCTGCAGGTGGG  
 CCGGACAAACCCCGCTCCGAGCCGCGCGGAGCGCCAGGGCACTTACCCTGAACCTTCCCCAGATCACCTTGTGGCAGCGCCCCCTGGTGTCCATCA  
 AGGTGGCGGCCAGATCAAGGAGGCCCTGTGACACCCGCGCCGACACCCGTGTGGAGGAGATCAACCTGCCCCGGCAAGTGAAGCCCAAGATGATC  
 GCGGCATCGGCGGCTTATCAAGGTGCGCCAGTACGACCCAGATCCTGATCGAGATCTGGGCAAGAAAGGCCATCGGCACCGTGTGGTGGGCCCCACCCC  
 CGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGTGCACTTGAACCTTCCCATCTCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCA  
 TGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTCGAGGAGATGGAGAAGGAGGCAAGATCACCAAG  
 ATCGGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCG  
 CACCCAGGACTTCTGGGAGGTGCACTGGGCATCCCCACCCCGCGGCTGAAGAAGAAGTCCGTGACCCGTGTGGACGTGGCGGACGCTTACTTCT  
 CCGTGCCCTGGACGAGGGCTTCGCAAGTACACCGCTTCACTTCCATCAACAACGAGACCCCGGCAATCCGCTACCACTACAACGTGCTGCC  
 CAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCTCCATGACCAAGATCCTGGAGCCCTTCCGCGCCAGAACCCCGAGATCGTGATCTACCACTACAT  
 GGACGACCTGTACGTGGGTCCGACCTGGAGATCGGCAGCACCCGCGCCAGATCGAGGAGCTGCGCGAGACCTGCTGAAGTGGGGCTTACCAACCCCG  
 ACAAGAAGCACCAAGAGGAGCCCCCTTCTGTGGATGGGTACGAGTGCACCCCGCAAGTGGACCGTGCAGCCCATCCAGCTGCCGAGAAGGACTCC  
 TGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCTCCAGATCTACCCCGCATCAAGTGGCGCAGCTGTGCAAGCTGTGCCGG  
 CGCAAGGCCCTGACCGACATCGTGCCCTGACCGAGGAGGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGAAGGAGCCCCGTGCACGGCGTGTACT  
 ACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCACGACCAAGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGC  
 AAGTACGCCAAGATGGCACCGCCACACCAACGACGTGAAGCAGTACCGAGGCCGTGCAGAAGATCGCCATGGAGTCCATCGTGATCTGGGGCAAGAC  
 CCCCAGTTCCGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCC  
 CCCCCCTGGTGAAGCTGTGTACAGCTGGAGAAGGAGCCCATCGCCGCGCGCGAGACCTTCTAGTGGACGGCGCGCCCAACCCGAGACCAAGATCGGC  
 AAGCCCGCTACGTGACCGACCGCGCGCGCGAGATCGTGTCCCTGACCGAGACCAACCAAGAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCA  
 GGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGTGGTGAACCCAGA  
 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACTGTCTGTGGTGCCCGCCCAAGGGCATCGGCGGCAACGAGAGGTGGACAAGCTGGTGTCTCTCC  
 GGCATCCGCAAGGTGCTTCTTGACGGCATCGACAAGGCCCGAGGAGGACGAGAAGTACCCTCAACTGGCGGCCATGGCCCTCCGAGTTCACCT  
 GCCCCCATCGTGGCCAAAGGAGATCGTGGCTCTCTGCGACAAGTGCAGCTGAAGGCGGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCTGGC  
 AGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGTGGCGGTGACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCGCCGAGACCCGGCCAG  
 GAGACCGCTACTTCACTCCTGAAGTGGCGGCGCTGGCCCGTGAAGGTGATCCACCGCAACGGCTCCAACCTCACCTCCCGCCGCGCTGAAGCCCGC  
 CTGCTGGTGGCGCGCATCCAGCAGGAGTTCGGCATCCCCACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAATCATCG  
 GCCAGGTGGCGGACCGGCGAGCACCTGAAGACCGCGTGCAGATGGCGGTGTTCACTCCACAACCTTCAAGCGCAAGGCGGATCGGCGGCTACTCCGCC  
 GCGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGTGCAGAAGCAGATCATCAAGTCCAGAATCTCCGCTGTACTACCGCGACTC  
 CCGGACCCCATCTGGAAGGGCCCCCGCAAGCTGTGTGAAGGGCGAGGGCGCGTGGTGTATCCAGGACAACCTCCGACATCAAGTGGTGGCCCCCGCGCA  
 AGGCCAAGATCATCCGGACTACGGCAAGCAGATGGCCGCGCGCGCTGGTGGCGCGCCGCGCAGGACGAGGACTAA

155/178

Fig. 113A

## 70. 2003 CON D pol. PEP

FFRENLAFFQKGAGELSSQTRANSRSLRVWGGDNPLSETGAERQGTVSFNFQITLWQRPVLTIKIGQLKEALLDTGADDTVLEEIN  
 LPGWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPKPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISRIGPENPYNTPIFAIAKKKDKSTKRWKLVDERELNKRQDFWEVQLGIPHPAGLKKKSVTLVDVGDAYFSVPLDE  
 DFRKYTAFTIPINNTPGIRYQYNVLPQGWKGSPIAQQSSMTKILEPFRKQNPFIYIYQMDLLYVGSDDLEIGQHRTKIEELREHLLRWGE  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPEKESWTVDIQLVGLKNWASQIYPGKVRQLCKLLRGTKALTEVIPLTEEAEELELAEN  
 REILKEPVHGVYDPSKDLIAEIQKQGGQWTYQIYQEPFNKLTGKYARMRGAHTNDVKOLTEAVQKIAIESIWIWGTPKFRLPIQKETW  
 ETWTEYWOATWIPWEFVNTPLVKLWYQLEKEPIIGAETFYVDGAANRETKLGKAGYVTDGRQKVPLTDTTNQKTELQAINLALQDSG  
 LEVNIIVDSQYALGIIQAOPDKSESELVSQIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVNSGIRKVLFLDGDIDKAQEEHEKYHNNWVRAM  
 ASDENLPPVVAKEIVASCDKCOLKGEAMHGQVDCSPGIWQLDCTHLEGVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVH  
 TDNGSNFTSAAVKAACWAGIKQEFGIPYNPQSQGVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIIAT  
 DIQTKELQKQIIKIQNFVYYRDSRDPWKGPAKLLWKGEAGAVIQDNSDIKVVPRRKVKIIRDYGKQMGAGDDCVASRQDED\$

156/178

Fig. 114A

## 71. 2003 CON F1 pol. PEP

FFRENLAFFQGEARKEFPSEQTRANSRSLRVQRGDNPLSEAGAERRGTVPSSLFPQITLWQRPVLTIKIGQLKEALLDTGADDTVLEDI  
 NLPGWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPKPGMDGPKVKQWPLTEE  
 KIKALTEICTEMEKEGKISRIGPENPYNTPVFAIAKKKDKSTKRWKLVDERELNKRQDFWEVQLGIPHPAGLKKKSVTLVDVGDAYFSVPLD  
 KDFRKYTAFTIPSVNNETPGIRYQYNVLPQGWKGSPIAQQSSMTKILEPFRKQNPFIYIYQMDLLYVGSDDLEIGQHRTKIEELREHLLKMG  
 FTTDPKKHQKEPPFLWMGYELHPDKWTVQPIQLPKDSWTVDIQLVGLKNWASQIYPGKVRQLCKLLRGAKALTDIVPLTAAEAELELAE  
 NREILKEPVHGVYDPSKDLIAEIQKQGGQWTYQIYQEPFNKLTGKYAKMRSHTNDVKOLTEAVQKIALESIVIWGTPKFRLPIQKET  
 WDTWWTDYWOATWIPWEFVNTPLVKLWYQLETEPIVGAETFYVDGASNRETCKGKAGYVTDGRQKVVSLETETTNQKAELOAIHLALQDS  
 GSEVNIIVDSQYALGIIQAOPDKSESELVNQIEQLIQKEKVYLSWVPAHKGIGGNEQVDKLVSAGIRKILFLDGDIDKAQEEHEKYHNNWRA  
 MASDENLPPVVAKEIVASCDKCOLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKII  
 HTDNGSNFTSAAVKAACWAGIQEFGIPYNPQSQGVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIIA  
 TDIQTRRELQKQITKIQNFVYYRDSRDPWKGPAKLLWKGEAGAVIQDENSEIKVVPRRKAKIIRDYGKQMGAGDDCVAGRQDED\$

## Fig. 113B

2003\_CON\_D pol:OPT

TTCTTCCGGAGAACCTGGCCCTTCCCCAGGGCAAGGCCGGGAGCTGTCTCTCGAGCAGACCCGCGCCAACTCCCCACCTCCCGGAGCTGCGGGTGTG  
GGCGGGGACAAACCCCTGTCCGAGACCGGCGCCGAGCGCCAGGGCACCGTGTCTTCAACTTCCCCAGATCACCTGTGGCAGCGCCCTTGGTGACCA  
TCAAGATCGGCGGCCAGCTGAAGGAGGCCCTGTGGACACCGGCCGACACCGTGTGGAGGAGATCAACCTGCCCGGCAAGTGAAGCCCAAGATG  
ATCGCGGCATCGCGGGCTTTCATCAAGGTGGCCAGTACGACCAAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGTGGTGGCCCCAC  
CCCCGTGAACATCATCGCGCCGCAACCTGTGACCCAGATCGGCTGACCCCTGAACCTTCCCCATCTCCCCATCGAGACCGTGTGGCGTGAAGCCCG  
GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAAGATCAAGGCCCTGACCGAGATCTGACCGAGATGGAGAAGGAGGCAAGATCTCC  
CGCATCGGCCCCGAGAACCCCTACAACACCCCCATCTTCGCCATCAAGAAGAAGGACTCCACCAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAACAA  
GGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGGCCCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGCGGACGCCCTACT  
TCTCCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCACTCCCTCCATCAACAAGAGACCCCGCATCCGCTACCACTACAACGTGTG  
CCCCAGGGCTGGAAGGCTCCCCCGCCATCTTCCAGTCTTCCATGACCAAGATCTTGGAGCCCTTCCGCAAGCAGAACCCCGAGATCGTGATCTACCAGTA  
CATGGAACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGACCGCACCAAGATCGAGGAGCTGCGGAGACCTGTGCTGGCTGGGGCTTCAACACCC  
CCGACAAGAAGCACGAGGAGGCCCTTCTGTGGATGGGTACGAGTGCACCCCGACAAGTGGACCGTGCAGCCCATCAAGCTGCCCGAGAAGGAG  
TCCTGGACCGTGACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCTTCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGTGCG  
CGCACCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCCGTGCACGGCGTGT  
ACTAGACCCCTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACC  
GGCAAGTACGCCCGCATCGCGCGGCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGAGTCCATCGTGATCTTGGGGCAA  
GACCCCAAGTTCGGCTGCCCTGCCATCCAGAGGAGACCTGGGAGACCTGGTGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACA  
CCCCCCCCCTGGTGAAGCTGTGTACAGCTGGAGAAGGAGCCCATCATCGGGCCCGAGACCTTCTACGTGGACGGCCCGCCCAACCGGAGACCAAGCTG  
GGCAAGGCCGCTACGTACCGACCGCGGCCCGCCAGAGGTGGTGGCCCTGACCGACACCAACCAAGAGACCGAGCTGGAGGCCATCAACCTGGCCCT  
GCAGGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAAGTCCGAGTCCGAGTGGTGTCCC  
AGATCATCGAGCAGTGTCAAGAAGGAGAGGTGTACCTGGCTGGCTGGCCCTGAGTACCGCCCTGAGGAGTCCAGGCCAGCCCGACAAAGTGGTGTCC  
AAGGCATCCGCAAGTGTCTTCTTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCAACAACACTGGCGGCCATGGCCTCCGACTTCAA  
CCTGCCCCCGTGGTGCCCAAGGAGATCGTGGCCCTCCTGCGACAAGTGGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCT  
GGCAGCTGACTGCACCCACCTGGAGGGCAAGGTGATCTTGTGGCCGTGACGTGGCCCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACCGGC  
CAGGAGACCGCTACTTCTCTGTGAAGTGGCCCGCCGTGGCCCGTGAAGTGGTGCACACCGACAACGGCTCCAACCTCACCTCCCGCCCGCTGAAGC  
CGCTGTGTGGTGGCCGGCATCAAGCAGGAGTTCGGCATCCCCATAACCCCAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
TCGGCCAGGTTCGGACAGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGGCTCGGGCGCTACTCC  
GCCGGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGTGCAGAAGCAGATCATCAAGATCCAGAACCTTCCGGTGTACTACCGGA  
CTCCCGGACCCCATCTGGAAGGGCCCGCAAGCTGTGTGGAAGGGCGAGGGCGCCGTGGTGTATCCAGGACAACCTCCGACATCAAGGTGTGCCCCCGC  
GCAAGGTGAAGATCATCCCGGACTACGGCAAGCAGATGGCCGGCGAGCTGCGTGGCCCTCCCGCCAGGACGAGGACTAA

157/178

Fig. 114B

2003\_CON\_F1\_pol.OPT

TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGCGGAGCCCGCAAGTTCCCTCCGAGCAGACCCGCGCCAACTCCCCCGCTCCCGCGAGCTGCGCGTGCA  
GGCGGCGACAACCCCTGTCCGAGCCGCGCGAGCGCGCGGCAACGTTGCCCTCCCTGTCTTCCCCAGATCAACCTGTGGCAGCGCCCCCTGGTGA  
CCATCAAGATCGGCGGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCGCGGACGACACCGTGTGGAGGACATCAACCTGCCCGGCAAGTGAAGCCCAAG  
ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGAAGCAGTACGACCAATCTGTATCGAGATCTGGGCGCACAGGCCATCGGCACCGTGTGTTGGGCCC  
CACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGATCGGCTGCAACCTGAACTTCCCCATCTCCCCATCGAGACCGTGGCGTGAAGCTGAAGC  
CCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAGGAGGCAAGATC  
TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAA  
CAAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCATCCCCACCCCGCGGCTGAAGAAGAAGTCCGTGACCCGTGTGGACGTGGCGGACGCCCT  
ACTTCTCCGTGCCCTGGACAAGGACTTCCGCAAGTACACCGCCTTCAACATCCCTCCGTGAACAACGAGACCCCGGCATCCGTACAGTACAACGTG  
CTGCCCCAGGGCTGGAAGGCTCCCCCGCATCTTCCAGTGTCCATGACCAAGATCTCTGGAGCCCTTCCGACCAAGAACCCCGACATCGTGATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCAGCACCAAGATCGAGGAGCTGCGGAGCACTGCTGAAGTGGGCTTCACCA  
CCCCGACAAGAAGCACCAAGAGAGCCCCCTTCTGTGGATGGCTACGAGCTGACCCCGACAAAGTGGACCGTGCAGCCCATCCAGCTGCCCGACAAG  
GACTCCTGGACCGTGAACGACATCCAGAAGCTGTGGCAAGTGAACCTGGGCTCCAGATCTACCCCGCATCAAGTGAAGCAGCTGTGCAAGCTGCT  
GCGCGCGCCAAAGCCCTGAACCGACATCGTGCCTTGAACCGCGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCGTGCACGGCG  
TGTACTACGACCCCTCCAGGACCTGATCGCCGAGATCCAGAAGCAGGCGCAGGCGCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAG  
ACCGGCAAGTACGCCAAGATGCGCTCCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCGTGCAGAAGATCGCCCTGGAGTCCATCGTGATCTGGGG  
CAAGACCCCAAGTTCGCGCTGCCATCTGAAGGAGACCTGGGACACCTGTTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGA  
ACACCCCCCTGGTGAAGCTGTGGTACCACTGGAGACCGAGCCCATCTGTGGCGCCGAGACCTTCTACGTGGACGGCGCTCCAAACCGCGAGACCAAG  
AAGGGCAAGCGCGCTACGTGACCGCGGCGCGCAGAAAGTGGTGTCCCTGACCGAGACCAACCAAGAGGCGGAGCTGCAGGCCATCCACCTGGC  
CCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCCAAGTACGCCCTGGGCATCATCCAGGCCCGCGGACAAAGTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGCAGCTGATCCAGAAGGAGAGTGTACCTGTCTGGTGGTGGCGGCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAAGTGGTG  
TCCGCGCGCATCCGCAAGATCCTGTTCTTGACGGCATCGACAAGGCCCAAGGAGGACGAGAAAGTACCAACAACCTGGCGGCCATGGCCTCCGACTT  
CAACCTGCCCCCGTGGTGGCCAAAGGAGATCGTGGCTCTCTGGACAAGTGGCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGGCGGTGCACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACC  
GGCAGGAGACCGCCTACTTCTATCTGAAGCTGGCGCGCGGTGGCCCGTGAAGATCATCCACACCGACAACGGCTCCAACCTTCACTCCGCGCGCGTGA  
GGCCGCTGTGTGGTGGCGCGCATCCAGCAGGAGTTCGGCATCCCTACAACCCCAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGAGCTGAAGAAG  
TCATCGGCCAGTGGCGACCGCCGAGACCTGAAGACCGCGTGCAGATGGCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGGCTATCGGCGGTAC  
TCCGCGCGGAGGCGATCATCGACATCATCGCCACCGACATCCAGACCGCGAGCTGCAGAAGCAGATCAACCAAGATCCAGAACCTTCCGCGTGTACTCCG  
CGACTCCCGCGACCCCGTGTGGAAAGGCGCGCGCAAGCTGCTGTGAAAGGCGAGGGCGCGGTGGTGTATCCAGGACAACCTCCGAGATCAAGTGGTGGCCCC  
GCGGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGCGCAGCTGCGTGGCGCGCGCCAGGACGAGGACTAA

158/178

Fig. 115A

## 72. 2003 CON F2 pol. PEP

FFRENLAFOQGEAREFSSEQTRANSPASRELVRRRGDSPLPEAGAERQGTGSSLDFFQITLWQRPVLTIKVGGQLREALLDTGADDTVLEEDI  
 NLPQKWKPMIGGIGGFIKVRQYDQIPIEICGKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEE  
 KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKRQDFWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVPLD  
 KEFRKYAFTIPISINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRANKNPEIYIYQYMDLTVGSDLEIGQHRTKIEELREHLLRWG  
 FTTDPKKHQKEPPFLWMGYELHPDKWTVQAIQLPDKSSWTVNDIQKLVGKLNWASQIYPGIRVKKHLCKLLRGAKALTDVVPVLTAAEAELELAE  
 NREILKEPVHGVYDPSKDLIAEIQKQCHDQWYQIYQEPHKNLKTGKYARRKSAHTNDVKQTEVVQKIATEGIVIGKVPKFRPLPIQKET  
 WEIWWTEYWQATWIPWEFEVNTTPLLVLWYQLETEPIVGAETFYVDGAANRETKLGKAGYVTDGRQKVPLTETTNOKTELQAIHLALQDS  
 GSEVNIIVTDSQYALGIIQAHDPKSESELVNOIIIEQLIQKERVYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
 MASDENLPPVVAKEIVASCDKQKGEAMHGQVDCSPGIWQLDCTHLEGGIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKII  
 HTDNGSNFTSTVVKAAACWWAGIQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIIA  
 TDIQTKELQKQITKIQNFRVYFRDSRDPVWKGPAPKLLWKGEAVVIQDNNEIKVVPRRKAKIIRDYGKQMGAGDDCVAGRQDED\$

159/178

Fig. 116A

## 73. 2003 CON G pol. PEP

FFRENLAFOQGEAREFSSEQTRANSPTRRELVRRRGDSPLPEAGAEGKGAISLSEFPQITLWQRPVLTIKVGGQLREALLDTGADDTVLEEDIN  
 LPQKWKPMIGGIGGFIKVRQYDQIILIEISGKAIGTVLVGPTPINIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKDKSTKWRKLVDFRELNKRQDFWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVPLDE  
 NFRKYAFTIPSTNNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRANKNPEIYIYQYMDLTVGSDLEIGQHRAKIEELREHLLRWG  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQAIQLPDKESWTVNDIQKLVGKLNWASQIYPGIRVKKHLCKLLRGAKALTDIVPLTAAEAELELAE  
 REILKEPVHGVYDPSKDLIAEVQKQGLDQWYQIYQEPYKNLKTGKYAKRGSHTNDVKQTEVVQKIATEGIVIGKVPKFRPLPIRKETW  
 EVWWTEYWQATWIPWEFEVNTTPLLVLWYRLTEPIPGAETYYVDGAANRETKLGKAGYVTDKQKQKIITLTTTNOKTELQAIHLALQDSG  
 SEVNIIVTDSQYALGIIQAQDRSESELVNOIIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRA  
 ASDENLPPVVAKEIVASCDKQKGEAMHGQVDCSPGIWQLDCTHLEGGIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH  
 TDNGSNFTSAVKAACWANITQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIIAS  
 DIQTKELQKQITKIQNFRVYFRDSRDPVWKGPAPKLLWKGEAVVIQDNNEIKVVPRRKAKIIRDYGKQMGAGDDCVAGRQDED\$







161/178

Fig. 116B

2003\_con\_g pol.1.OPT

TTCTTCCGGAGAACTGGCCCTCCAGCAGGGCGAGGCCCGGAGTTCTCTCCGAGCAGGCGCGGCCCAACTCCCCACCCCGCGCGAGCTGCGCGTGC  
 CCGCGGAGTCCCCCTGCCCCGAGCGCGCGAGGGCAAGGGGCCATCTCCCTGTCTTCCCCAGATCACTCCCTGTGGCAGCGCCCCCTGGTGACCG  
 TGAAGATCGCGCGCAGCTGATCGAGGCCCTGCTGGACACCGCGCGCAGACACCGTGTGGAGGAGATCAACTGCCCCGGCAAGTGAAGCCCAAGATG  
 ATCGCGGCATCGCGCGCTTCAATCAAGTGGCCAGTACGACCCAGATCTGATCGAGATCTCCGGCAAGAGGCCATCGCATCGCACCGTGTGGTGGCCCCAC  
 CCCCATAACATCATCGGCGCAACATGCTGACCCAGATCGGCTGCACCTGAACTTCCCCATCTCCCCATCGAGACCGTGCCTGAGAGGAGGCAAGATCTCC  
 GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGACCGAGATCGACCGAGATGGAGGAGCTCCCGGAGCTGAACAA  
 AAGATCGGCCCCGAGAACCCCTACAACACCCCCCATTTCCGCTATCAAGAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAACAA  
 GCGACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCCCACCCCGCGGCTGAAGAAGAAGTCCGTGACCCGTGTGGACGTGGCGGACGCGCTACT  
 TCTCCGTGCCCTGGACGAGAACTTCCGCAAGTACACCGCTTCACTCCCTCCACCAAGATCTGGAGCCCTTCCGACCAAGAACCCCGAGATCGTGATCTACCACTA  
 CCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCAGTCTCCATGACCAAGATCTGGAGCCCTTCCGACCAAGAACCCCGAGATCGTGCTGGGCTTCAACACC  
 CATGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGTCCGCGAGACCTGCGGAGACCTGCTGGGCTTCAACACC  
 CCGACAAGAACCAAGAGGAGCGCCCTTCTGTGGATGGGTACGAGTGCACCCCGACAAGTGGACCTGTGAGCCCCATCCAGCTGCCGACAAAGGAG  
 TCCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCTCCAGATCTACCCGGCATCAAGTGAAGCAGCTGTGCAAGCTGTGCG  
 CGGCGCAAGGCCCTGACCGACATCGTGCCCTGACCCCGAGGCGGAGCTGGAGCTGGCCGAGAACCGGAGATCTTGAAGGAGCCCCGTGCACGGGCTGT  
 ACTACGACCCCTCCAAGGAGCTGATCGCCGAGTGCAGAACGAGGCGCTGGACCATGACCTGACCTACAGATCTACAGGAGCCCTACAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGCGGGCTCCGCCACACCAACGACGTGAAGCAGCTGACCGAGGTGTGCAAGATCGCCACCGAGTCCATCGTGATCTGGGGCAA  
 GACCCCAAGTCAAGCTGCCCATCCGCAAGGAGACCTGGGAGGTGTGGTGGACCGATCTGGAGGCCACCTGAGTGGACGGCGCCGCAACCGGAGACCAAGCTG  
 CCCCCCTGTGTGTAAGCTGTGTACCGCTGGAGACCGAGCCCATCCCCGGCGCGGAGACCTACTAGTGGACGGCGCGAGCTGCAGGCCATCCACCTGGCCCT  
 GGCAAGGCCGGTACGTGACCGACAAGGCAAGCAGAGATCATACCTGACCGAGACCAACCAAGAGGCCGAGCTGCAGGCCATCCACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACCGCTGGGCATCATCCAGGCCAGCCCGACCGCTCCGAGTCCGAGTGGTGAACC  
 AGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGTCTGGTGCCCGCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTGTCC  
 TCCGGCATCCGCAAGGTGTCTCTGACGGCATGCACAGGCCCAGGAGGAGCACGAGCGCTACCACTCCAAGTGGCGGCCATGGCCTCCGACTTCAA  
 CCTGCCCCCATCGTGGCCAAAGGATCGTGGCTCTCTGGACAAGTGGCAGCTGAAGGGGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGACATCT  
 GGAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGGTGGCCGTGCACGTGGCTTACATCGAGGCCAGGTGATCCCCCGCGAGACCGGC  
 CAGGAGACCGCTACTTCACTCTGAAGTGGCGCGCGTGGCCCTACACCCCAAGTCCAGACCGACAACGGCTCCAACCTCCGCGCGCTGAAGGC  
 CGCTGTGTGGGCCAACATCACCCAGGAGTTCGGCATCCCCCTACAACCCCAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCGACCCAGGCCGAGCACCTGAAGACCGCGTGCAGATGGCCGTGTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGGCGGCTACTCC  
 GCGGCGAGGCGCATATCGACATCATCGCTCCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCCAGAACTTCCGCTGTACTACCGGA  
 CTCCCGGACCCCATCTGGAAGGGCCCCCAAGCTGCTGTGGAAGGGCGAGGCGCGCTGGTGTATCCAGGACAACAACGAGATCAAGGTGGTGTCCCCCGC  
 GCAAGGCCAAGATCATCCCGGACTACGGCAAGCAGATGGCCGGCGGCGGCTGGCTGGCCCGCGGACGAGGACTAA

Fig. 117A

74. 2003 CON H pol. PEP  
 FFRENLAFOQREARKFSPEQARANSPTSRELVRRGDDPLSEAGAEQGTSFPQITLWQRPLVTVKIEGQREALLDTGADDTVLEEINL  
 PGKWKPKMIGGIGGFIKVRQYEQVAIEICGKKAIGTVLVGPTPVNIIGRNILTIQIGCTLNFPIETVPVKLPGMDGPKVKQWPLTEEKI  
 KALTEICIEMEKEGKISKIGPENPYNTPIFAIKKDDSTKWRKLVDFRELNKRQTQDFWEVQLGIPHPAGLKKKSVSLDVGDAYFSVPLDKD  
 FRKYTAFTIP SINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPENIIYQYMDLIVGSDLEIGQHRAKIEELRAHLLRWGFT  
 TPDKKHQKEPPFLWMGYELHPDKWTVPVKLPEKDSWTVNDIQKLVGKLNWASQIYPGKVKQCKLRLGAKALTDIVPLTKEAELELAENR  
 EILREPVHGYYDPSKDLIAEIQKQGDQWYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIATESIVIWGKIPKFRLP IQKETWE  
 TWTEHWQATWIPWEFVNTPHLVKLWYQLETEPIAGAEYIYVDGAANRETKIGKAGYVTDGKQKVSLTETTNQKTELQAIYLAQDSGL  
 EVNIVTDSQYALGIIQAQPKSESELVNQIIIEELIKKEKYL SWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHRYHNNWRAMA  
 SDFNLPIVAKELVASCDCQKLGAEAMHGQVDCSPGIWQLDCTHLEGKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKMIHT  
 DNGSNFTSAAVKAACWADIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLRTAVQMAVFIHNEKRKGGIGGYSAGERIIDIIATD  
 IQTKELQKQISKIQFRVYRDSRDPINWGPALLWKGEAVVIQDNSEIKVVPRRKAKIIRDYGMAGDDCVAGRQDED\$

162/178

Fig. 118A

75. 2003 CON 01 AE pol. PEP  
 FFRENLAFOQKAGFESSEQTRANSPSTRKLGDDGRDNLLEAGAEQGTSSSFSPQITLWQRPLVTVKIGGQKEALLDTGADDTVLEDI  
 NLPKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTIQIGCTLNFPIETVPVTLLKPGMDGPKVKQWPLTEE  
 KIKALTEICKEMEEEGKISKIGPENPYNTPVFAIKKDDSTKWRKLVDFRELNKRQTQDFWEVQLGIPHPAGLKKKSVTLVDVGDAYFSVPLD  
 ESFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPENIIYQYMDLIVGSDLEIGQHRAKIEELRAHLLSWG  
 FTTDPKKHQKEPPFLWMGYELHPDRWTVPVKLPEKDSWTVNDIQKLVGKLNWASQIYPGKVKQCKLRLGAKALTDIVPLTKEAELELAE  
 NREILKTPVHGYYDPSKDLVAVQKQGDQWYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIATESIVIWGKIPKFRLP IQRET  
 WETWMEYWQATWIPWEFVNTPHLVKLWYQLETEPIAGAEYIYVDGAANRETKIGKAGYVTDGKQKVSLTETTNQKTELHAIHLALQDS  
 GSEVNIIVTDSQYALGIIQAQPKSESELVNQIIIEELIKKEKYL SWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHRYHNSNWR  
 MASDFNLPIVAKELVASCDCQKLGAEAMHGQVDCSPGIWQLDCTHLEGKVLVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI  
 HTDNGSNFTSAAVKAACWANVRQEFIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLRTAVQMAVFIHNEKRKGGIGGYSAGERIIDIIA  
 TDIQTKELQKQITKIQFRVYRDSRDPINWGPALLWKGEAVVIQDNSEIKVVPRRKAKIIRDYGMAGDDCVAGRQDED\$

Fig. 117B

2003\_CON\_H\_pol.OPT

TTCTTCCGCGAGAACCTGGCCCTCCAGACGCGGAGGGCCCGCAAGTTCTCCCCCGAGCAGGCCCCGCGCAACTCCCCACCTCCCCGCGAGCTGCGCGTGCG  
 CCGCGCGACGACCCCTGTCCGAGGCGGCGCGAGGGCCAGGGCACTCCCTGTCTTCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGACCGTGA  
 AGATCGAGGGCCAGCTGCGCGAGCCCTGTGGACACCGCGCGACGACCCGTGTGGAGGAGATCAACCTGCCCGCAAGTGAAGCCCAAGATGATC  
 GCGGCATCGCGGCTTCATCAAGGTGCGCCAGTACGAGCAGGTGGCCATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTGTGGTGGCCCCACCCC  
 CGTGAACATCATCGGCGGCAACATCCTGACCCAGATCGGCTGCACTTCCCATCTCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGGCA  
 TGGACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCATCGAGATGGAGAAGGAGGCAAGATCTCCAAG  
 ATCGGCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAAGAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAGCTGAACAAGCG  
 CACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCACCCCGCCCTGAAGAAGAAGTCCGTGTCCGTGTGGACGTGGCGGACGCCCTACTTCT  
 CCGTGCCCTTGGACAAGGACTTCGGCAAGTACACCGCTTCAACATCCCTCCATCAACAACGAGACCCCGGCATCCGCTACAGTACAACAGTGTGCCC  
 CAGGGCTGGAAGGCTCCCCCGCATCTTCCAGTCTTCCATGACCAAGATCCTGGAGCCCTTCGCAAGCAGAACCCCGAGATGATCATACCAAGTACAT  
 GGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCCCAAGATCGAGGAGCTGCGCGCCACCTGTGGCTGGGGCTTCAACACCCCG  
 ACAAGAAGCACCAAGAGGCCCTTCTGTGGATGGCTACGAGTGCACCCCGCAAGTGGACCGTGCAGCCCGTGAAGCTGCCCGAGAGGACTCC  
 TGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAAGTGGGCTCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGTGCGGG  
 CGCAAGGCCCTGACCGACATCGTGGCCCTGACCAAGGAGGCGGAGTGGAGCTGGCCGAGAACCGCGAGATCTTGGCGGAGCCCGTGCACCGCGTGTACT  
 ACGACCCCTTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCCCGACAGTGGACCTACCAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGC  
 AAGTACGCCAAGATCGCACCGCCACACCAACGACCTGAAGCAGTGAAGGCGCTGCAGAAAGTCCGCCACCGAGTCCATCGTGTATCTGGGGCAAGAT  
 CCCAAGTTCGCGCTGCCATCCAGAAGGAGACCTGGTGGAGACCTGGTGAACCGAGACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCC  
 CCCACCTGGTGAAGCTGTGGTACCAGCTGGAGACCGAGCCCATCGCCGGCGCGAGACCTACTACGTGGACGGCGCGCCCAACCGCGAGACCAAGATCGGC  
 AAGGCCGCTACGTGACCGACCGCGGCAAGCAGAAGTGTGTCCCTGACCGAGACCAACCAAGACCGAGCTGCAGGCCATCTACCTGGCCCTGCA  
 GGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGTCCGAGTGGTGAACCCAGA  
 TCATCGAGGAGCTGATCAAGAAAGGAGAGGTGTACCTGTCTGGTGCCCGCCACAAGGGCATCGGCGCAACGAGCAGGTGGACAAGCTGGTGTCTCTCC  
 GGATCCGCAAGGTGTCTTCTGGACGGCATCGACAAGGCCCGAGGAGCACGAGCGCTACCAACAACACTGGCGGCCATGGCTCCGACTTCAACCT  
 GCGCCCATCGTGGCCAAAGAGATCGTGGCCCTCTGCGACAAGTGCAGCTGAAGGGGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCTGGC  
 AGCTGGACTGCACCCACCTGAGGGCAAGGTGATCTGTGGTGGCCGTGCACGTGGCTCCGGTACATCGAGGCCGAGGTGATCCCCGCGAGACCGGCCAG  
 GAGACCGCTACTTCATCTGAAGCTGGCCGCGCTGGCCCGTGAAGATGATCCACACCGACACCGGCTCCAACCTCACTCCGCGCGCTGAAGGCCG  
 CTGCTGTGGCGGACATCCAGCAGGATTCGGCATCCCCACAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCATCG  
 GCCAGGTGGCGGACCGAGACCTGGCGACCGCGTGCAGATGGCCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGGCATCGGCGCTACTCCGCC  
 GCGAGCGCATCATCGACATCATCGCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCTCCAAGATCCAGAAGTTCGCGGTGTACTACCGCGACTC  
 CCGGACCCCATCTGGAAGGGCCCCGCAAGCTGTGTGGAGGGCGAGGCCGTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGGCCCCGCCGA  
 AGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCCGCGACGACTGCGTGGCCGCGCGCAGGACGAGGACTAA

163/178

Fig. 118B

2003 CON 01 AE pol.OPT

[illegible]

165/178

Fig. 119A

76. 2003 CON 02 AG pol.PEP  
 FFRENLA~~FQ~~GE~~AR~~K~~F~~SE~~Q~~TG~~T~~NS~~P~~TS~~R~~EL~~W~~DG~~G~~R~~N~~L~~L~~SE~~A~~GE~~G~~Q~~T~~ISS~~F~~N~~F~~P~~Q~~IT~~L~~W~~Q~~R~~P~~L~~V~~T~~V~~R~~I~~GG~~Q~~L~~E~~AL~~L~~D~~T~~G~~A~~D~~D~~T~~V~~LE~~E~~I  
 NL~~P~~G~~K~~W~~K~~P~~K~~M~~I~~G~~G~~I~~G~~F~~I~~K~~V~~R~~Q~~D~~Q~~IL~~E~~IC~~G~~K~~A~~I~~G~~T~~V~~L~~V~~G~~P~~T~~P~~V~~N~~I~~I~~GR~~N~~M~~L~~T~~Q~~IG~~C~~T~~L~~N~~F~~P~~I~~SP~~I~~ET~~V~~P~~V~~L~~K~~P~~G~~M~~D~~G~~P~~K~~V~~K~~Q~~W~~P~~L~~T~~EE  
 K~~I~~K~~A~~L~~T~~D~~I~~C~~T~~EME~~K~~E~~G~~K~~I~~S~~K~~I~~G~~P~~E~~N~~P~~Y~~N~~T~~P~~V~~F~~A~~I~~K~~K~~D~~S~~T~~K~~W~~R~~K~~L~~V~~D~~E~~R~~E~~L~~N~~K~~R~~T~~Q~~D~~F~~E~~V~~Q~~L~~G~~I~~P~~H~~P~~A~~G~~L~~K~~K~~K~~S~~V~~T~~V~~D~~V~~G~~D~~A~~Y~~F~~S~~V~~P~~L~~D~~  
 K~~D~~F~~R~~K~~Y~~T~~A~~F~~T~~I~~P~~S~~V~~N~~N~~E~~T~~P~~G~~I~~R~~Y~~Q~~Y~~N~~V~~L~~P~~Q~~G~~W~~K~~G~~S~~P~~A~~I~~F~~Q~~A~~S~~M~~T~~K~~I~~L~~E~~P~~F~~R~~T~~K~~N~~P~~E~~I~~V~~I~~Y~~Q~~M~~D~~D~~L~~Y~~V~~G~~S~~D~~L~~E~~I~~G~~H~~R~~A~~K~~I~~E~~E~~L~~R~~E~~H~~L~~L~~R~~W~~G~~  
 F~~T~~T~~P~~D~~K~~K~~H~~Q~~E~~P~~P~~L~~M~~G~~Y~~E~~L~~H~~P~~D~~K~~W~~T~~V~~Q~~I~~Q~~L~~P~~E~~K~~D~~S~~W~~T~~V~~N~~D~~I~~O~~K~~L~~V~~G~~K~~L~~N~~W~~A~~S~~Q~~I~~Y~~A~~G~~I~~K~~V~~K~~Q~~L~~C~~K~~L~~L~~R~~G~~A~~K~~A~~L~~T~~D~~I~~V~~T~~L~~T~~E~~E~~A~~E~~L~~E~~L~~A~~E~~  
 N~~R~~E~~I~~L~~K~~E~~P~~V~~H~~G~~V~~Y~~Y~~D~~P~~T~~K~~D~~L~~I~~A~~E~~I~~Q~~K~~Q~~D~~Q~~W~~T~~Y~~Q~~I~~Y~~Q~~E~~P~~F~~K~~N~~L~~K~~T~~G~~K~~Y~~A~~K~~M~~R~~S~~A~~H~~T~~N~~D~~V~~K~~L~~T~~E~~V~~V~~Q~~V~~A~~T~~E~~S~~I~~V~~I~~W~~G~~K~~T~~P~~K~~F~~R~~L~~P~~I~~Q~~R~~E~~T~~  
 W~~E~~A~~W~~M~~E~~Y~~N~~Q~~A~~T~~W~~I~~P~~E~~W~~E~~F~~V~~N~~T~~P~~L~~V~~K~~L~~W~~Y~~Q~~L~~E~~K~~D~~P~~I~~V~~G~~A~~E~~T~~F~~Y~~V~~D~~G~~A~~A~~N~~R~~E~~T~~K~~L~~G~~A~~G~~Y~~V~~T~~D~~R~~G~~R~~Q~~K~~V~~V~~S~~L~~T~~E~~T~~T~~N~~Q~~K~~T~~E~~L~~H~~A~~I~~H~~L~~A~~L~~Q~~D~~S  
 G~~S~~E~~V~~N~~I~~V~~T~~D~~S~~Q~~Y~~A~~L~~G~~I~~I~~Q~~A~~Q~~P~~D~~R~~S~~E~~S~~E~~L~~V~~N~~Q~~I~~I~~E~~K~~L~~I~~E~~K~~D~~K~~V~~Y~~L~~S~~W~~V~~P~~A~~H~~K~~G~~I~~G~~G~~N~~E~~Q~~V~~D~~K~~L~~V~~S~~N~~G~~I~~R~~K~~V~~L~~F~~D~~I~~D~~G~~I~~D~~K~~A~~Q~~E~~E~~H~~E~~R~~Y~~H~~S~~N~~W~~R~~A  
 M~~A~~S~~D~~F~~N~~L~~P~~P~~I~~V~~A~~K~~E~~I~~V~~A~~S~~C~~D~~K~~Q~~L~~K~~G~~E~~A~~M~~H~~G~~Q~~V~~D~~C~~S~~P~~G~~I~~W~~Q~~L~~D~~C~~T~~H~~L~~E~~G~~K~~I~~I~~L~~V~~A~~V~~H~~V~~A~~S~~G~~Y~~I~~E~~A~~E~~V~~I~~P~~A~~E~~T~~G~~Q~~E~~T~~A~~Y~~F~~I~~L~~K~~L~~A~~G~~R~~W~~P~~V~~K~~I~~  
 H~~T~~D~~N~~G~~S~~N~~F~~T~~S~~A~~A~~V~~K~~A~~C~~W~~A~~N~~V~~T~~Q~~E~~F~~G~~I~~P~~Y~~N~~P~~Q~~S~~Q~~V~~V~~E~~S~~M~~N~~K~~E~~L~~K~~K~~I~~I~~Q~~V~~R~~D~~Q~~A~~E~~H~~L~~K~~T~~A~~V~~Q~~M~~A~~V~~F~~I~~H~~N~~F~~K~~R~~K~~G~~G~~I~~G~~G~~Y~~S~~A~~G~~E~~R~~I~~I~~D~~I~~A  
 S~~D~~I~~Q~~T~~K~~E~~L~~Q~~K~~I~~T~~K~~I~~Q~~N~~F~~R~~V~~Y~~R~~D~~S~~R~~D~~P~~I~~W~~K~~G~~P~~A~~K~~L~~L~~W~~K~~G~~E~~G~~A~~V~~V~~I~~Q~~D~~N~~S~~D~~I~~K~~V~~V~~P~~R~~R~~K~~A~~K~~I~~I~~R~~D~~Y~~G~~K~~Q~~M~~A~~G~~D~~D~~C~~V~~A~~G~~R~~Q~~D~~E~~D~~§~~

Fig. 120A

77. 2003 CON 03 AB pol.PEP  
 FFRENLA~~FQ~~RE~~AR~~K~~F~~SE~~Q~~T~~R~~AIS~~P~~TS~~R~~K~~L~~W~~D~~G~~G~~R~~D~~N~~L~~P~~E~~T~~G~~T~~E~~R~~Q~~T~~A~~SS~~F~~N~~F~~P~~Q~~IT~~L~~W~~Q~~R~~P~~L~~V~~T~~V~~R~~I~~GG~~Q~~L~~E~~AL~~L~~D~~T~~G~~A~~D~~D~~T~~V~~LE~~E~~I  
 NL~~P~~G~~K~~W~~K~~P~~K~~M~~I~~G~~G~~I~~G~~F~~I~~K~~V~~R~~Q~~D~~Q~~IL~~E~~IC~~G~~K~~A~~I~~G~~T~~V~~L~~V~~G~~P~~T~~P~~V~~N~~I~~I~~GR~~N~~M~~L~~T~~Q~~L~~G~~C~~T~~L~~N~~F~~P~~I~~SP~~I~~ET~~V~~P~~V~~T~~L~~K~~P~~G~~M~~D~~G~~P~~K~~V~~K~~Q~~W~~P~~L~~T~~EE  
 K~~I~~K~~A~~L~~T~~D~~I~~C~~K~~EME~~K~~E~~G~~K~~I~~S~~K~~I~~G~~P~~E~~N~~P~~Y~~N~~T~~P~~V~~F~~A~~I~~K~~K~~D~~S~~T~~K~~W~~R~~K~~L~~V~~D~~E~~R~~E~~L~~N~~K~~R~~T~~Q~~D~~F~~E~~V~~Q~~L~~G~~I~~P~~H~~P~~A~~G~~L~~K~~K~~K~~S~~V~~T~~V~~D~~V~~G~~D~~A~~Y~~F~~S~~V~~P~~L~~D~~  
 Q~~D~~F~~R~~K~~Y~~T~~A~~F~~T~~I~~P~~S~~T~~N~~N~~E~~T~~P~~G~~I~~R~~Y~~Q~~Y~~N~~V~~L~~P~~Q~~G~~W~~K~~G~~S~~P~~A~~I~~F~~Q~~S~~S~~M~~T~~K~~I~~L~~E~~P~~F~~R~~K~~Q~~N~~P~~E~~I~~V~~I~~Y~~Q~~M~~D~~D~~L~~Y~~V~~G~~S~~D~~L~~E~~I~~G~~H~~R~~T~~K~~I~~E~~E~~L~~R~~E~~H~~L~~L~~R~~W~~G~~  
 F~~T~~T~~P~~D~~K~~K~~H~~Q~~E~~P~~P~~L~~M~~G~~Y~~E~~L~~H~~P~~D~~K~~W~~T~~V~~Q~~I~~Q~~L~~P~~E~~K~~D~~S~~W~~T~~V~~N~~D~~I~~Q~~K~~L~~V~~G~~K~~L~~N~~W~~A~~S~~Q~~I~~Y~~A~~G~~I~~K~~V~~R~~Q~~L~~C~~K~~L~~L~~R~~G~~A~~K~~A~~L~~T~~E~~V~~I~~P~~L~~T~~A~~E~~A~~E~~L~~E~~L~~A~~E~~  
 N~~R~~E~~I~~L~~K~~E~~P~~V~~H~~G~~V~~Y~~Y~~D~~P~~S~~K~~D~~L~~V~~A~~E~~I~~Q~~K~~Q~~G~~Q~~W~~T~~Y~~Q~~I~~Y~~Q~~E~~P~~F~~K~~N~~L~~K~~T~~G~~K~~Y~~A~~R~~L~~R~~G~~A~~H~~T~~N~~D~~V~~K~~L~~T~~E~~A~~V~~Q~~K~~I~~A~~T~~E~~S~~I~~V~~I~~W~~G~~K~~T~~P~~K~~F~~K~~L~~P~~I~~Q~~K~~E~~T  
 W~~E~~T~~W~~W~~T~~E~~Y~~W~~Q~~A~~T~~W~~I~~P~~E~~W~~E~~F~~V~~N~~T~~P~~P~~L~~V~~K~~L~~W~~Y~~Q~~L~~E~~K~~E~~P~~I~~V~~G~~A~~E~~T~~F~~Y~~V~~D~~G~~A~~A~~N~~R~~E~~T~~K~~S~~G~~K~~A~~G~~Y~~V~~T~~D~~R~~G~~R~~Q~~K~~V~~V~~S~~L~~T~~D~~T~~T~~N~~Q~~K~~T~~E~~L~~Q~~A~~I~~H~~L~~A~~L~~Q~~D~~S~~  
 G~~L~~E~~V~~N~~I~~V~~T~~D~~S~~Q~~Y~~A~~L~~G~~I~~I~~Q~~A~~Q~~P~~D~~K~~S~~E~~S~~E~~L~~V~~S~~Q~~I~~I~~E~~Q~~L~~I~~K~~K~~E~~K~~V~~Y~~L~~A~~W~~V~~P~~A~~H~~K~~G~~I~~G~~G~~N~~E~~Q~~V~~D~~K~~L~~V~~S~~A~~G~~I~~R~~K~~V~~L~~F~~D~~G~~I~~D~~K~~A~~Q~~E~~A~~H~~E~~K~~Y~~H~~S~~N~~W~~R~~A  
 M~~A~~S~~D~~F~~N~~L~~P~~P~~V~~V~~A~~K~~E~~I~~V~~A~~S~~C~~D~~K~~Q~~L~~K~~G~~E~~A~~M~~H~~G~~Q~~V~~D~~C~~S~~P~~G~~I~~W~~Q~~L~~D~~C~~T~~H~~L~~E~~G~~K~~I~~I~~L~~V~~A~~V~~H~~V~~A~~S~~G~~Y~~I~~E~~A~~E~~V~~I~~P~~A~~E~~T~~G~~Q~~E~~T~~A~~Y~~F~~V~~L~~K~~L~~A~~G~~R~~W~~P~~V~~K~~I~~  
 H~~T~~D~~N~~G~~S~~N~~F~~I~~S~~T~~A~~V~~K~~A~~C~~W~~A~~G~~I~~K~~Q~~E~~F~~G~~I~~P~~Y~~N~~P~~Q~~S~~Q~~V~~V~~E~~S~~M~~N~~K~~Q~~L~~K~~Q~~I~~I~~Q~~V~~R~~D~~Q~~A~~E~~H~~L~~K~~T~~A~~V~~Q~~M~~A~~V~~F~~I~~H~~N~~F~~K~~R~~K~~G~~G~~I~~G~~G~~Y~~S~~A~~G~~E~~R~~I~~I~~D~~I~~A  
 T~~D~~I~~Q~~T~~K~~E~~L~~Q~~K~~Q~~I~~I~~K~~I~~Q~~N~~F~~R~~V~~Y~~R~~D~~S~~R~~D~~P~~I~~W~~K~~G~~P~~A~~K~~L~~L~~W~~K~~G~~E~~G~~A~~V~~V~~I~~Q~~D~~N~~D~~I~~K~~V~~V~~P~~R~~R~~K~~A~~K~~I~~I~~R~~D~~Y~~G~~K~~Q~~M~~A~~G~~D~~D~~C~~V~~A~~S~~R~~Q~~D~~E~~D~~§~~

**Fig. 119B**

2003 CON 02 AG pol.OPT

TTCTTCCGCGAAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCAAGTTCTCTCCGAGCAGACCGGCACCAACTCCCCCACTCCCCCGGAGCTGTGGACCG  
CGCCCGGACAACCTGCTGTCCGAGGCCGACCGAGGGCCAGGCACCATCTCTCTTCAACTTCCCCCAGATCACCTTGTGGCAGCGCCCCCTGGTGA  
CCGTGCGCATCGCGGCCAGCTGATCGAGGCCCTGCTGGACACCGGGCCGAGACACCGTGTGAGGAGATCAACTGCCCGGCAAGTGAAGCCCAAG  
ATGATCGGCGGCATCGGCGGCTTTCATCAAGGTGGCCAGTACGACCAGATCTCTGATCGAGATCTGGGAAGAGGCCATCGGCACCGTGTCTGTGGGCC  
CACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACCTGAACCTTCCCATCTCCCATCGAGACCGTGCCTGAAGCTGAAGC  
CCGGCATGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGCACCGAGATGGAGAAGGAGGCAAGATC  
TCCAAGATCGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAA  
CAAGCGCACCCAGGACTTCTGGGAGTGCAGTGGGCATCCCCACCCCGCGGCTGAAGAAGAAGATCCGTGACCGTGTGGACGTGGCGGACGCGCT  
ACTTCTCGTGCCTGGACAAGGACTTCCGCAAGTACACCGCTTCAACCATCCCTCCGTGAACAACGAGACCCCGGCATCCGTACCAGTACAACGTG  
CTGCCCCAGGGCTGGAAGGGCTCCCCGCCATCTTCCAGGCTCCATGACCAAGATCTTGAGCCCTTCCGACCAAGAACCCCGAGATCGTGATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGCTGCGGAGCACCTGCTGGCTGGGGCTTCAACA  
CCCCGACAAGAAGCACAGAGGCCCTTCTCTGTGGATGGCTACGAGCTGCACCCGACAAGTGGACCGTGCAGCCCCATCCAGCTGCCCGAGAAG  
GACTCTTGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCTCCAGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCT  
GGCGGGCCCAAGGCCCTGACCGACATCGTGACCTTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCCCGAGATCTCTGAAGGAGCCCCGTGCACGGG  
TGTACTACGACCCCAAGGACTGTATCGCCGAGATCCAGAAGCAGGGCCAGGACCAAGTGGACCTACCAGATCTACAGGAGCCCTTCAAGAACCTGAAG  
ACCGCAAGTACGCCAAGATCGCTCCGCCACACCAACGACTGAAGCAGCTGACCGAGGTGGTGCAGAAGGTGGCCACCGAGTCCATCGTGATCTGGGG  
CAAGACCCCAAGTTCGCTGCCCATCCAGCGGAGACCTGGGAGGCCCTGGTGGATGGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGA  
ACACCCCCCTGTTGAAGCTGTGTACAGCTGGAGAAGGACCCCATCGTGGGCCCGAGACCTTCTACGTGACCGCGGCCGCCAACCCCGAGACCAAG  
CTGGCAAGCCGGTACGTGACCGCGGCCGCCAGAAGTGGTGTCCCTGACCGAGACCAACCAAGACCGAGCTGCACGCCATCCACCTGGC  
CCTGCAGGACTCCGGCTCCGAGGTGAACATCTGTACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCCGACCGCTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGAAGCTGATCGAGAAGGACAAGTGTACCTGTCTTGGTGGTGGCCGCCACCAAGGGCATCGCGGCAACGAGCAGTGGACAAGCTGGTG  
TCCAACGGCATCCGCAAGTGTCTTGACGGCATCGACAAGCCCGAGGAGCAGCGCTACCACTCCAATGGCGGCCCATGGCCCTCCGACTT  
CAACCTGCCCCCATCGTGGCCAAGGATCGTGGCTCTTCGACAAAGTGGCAGTGAAGGGCAGGCCATGCACGGCCAGGTGGAATGCTCCCCCGGCA  
TCTTGGCAGTGGACTGCACCCACTGGAGGGCAAGATCATCTGGTGGCCGTGCACGTGGCCCTCCGGCTACATCGAGGCCGAGTGTATCCCCCGCCGAGACC  
GGCCAGGAGACCGCTACTTTCATCTGAAGCTGGCCGCCGTGGCCGTGAAGTGTATCCACACCGACAACGGCTCCAATTCACCTCCGCCGCCGTGA  
GGCCGCCCTGCTGTTGGCCAACTGACCCAGGATTCGGCATCCCCTACAACCCCGAGTCCAGGGCGTGGTGGAGTCCATGAACAAGAGCTGAAGAAG  
TTCATCGGCCAGGTGCGGACCAAGCCCGCTGCAGATGGCCGTGTTCATCCACAACCTTCAAGCGCAAGGGCGGCAATCGGGCGCTAC  
TCCGCCGGCAGCGCATCATCGACATCATCGCATCCAGACCAAGGAGCTGCAGAAGCAGATCAACCAAGATCCAGAATTCGCGTGTACTACCG  
CGACTCCCCGACCCCATCTGGAAGGCCCCCGCAAGCTGCTGTGAAGGGCGAGGGCGCGCTGGTGTATCCAGGACAACCTCCGACATCAAGGTGGTGGCCC  
CGCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCAGCACTGCGTGGCCGCCCGCCAGGACGAGGACTAA



**Fig. 120B**

2003 CON 03 AB pol.OPT

2003\_CON\_03\_AB\_pos1.OP1

TTCTTCGGGAGAACTTGGCCCTCCAGCAGCGGAGGCCCGCAAGTTCTCTCCGAGCAGACCCCGGCCCATCTCCCCACCTCCCGCAAGCTGTGGACGG  
CGGCCGGACAACCCCTGCCGAGACCGGCAACGAGGCCAGGGCAACGGCTCCTCTTCAACTTCCCCAGATCACTTGTGGCAGCGCCCCCTGGTGA  
CCGTGGCATCGCGGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCCGACACACCGTCTGGAGACATCAACTGCCCCGCAAGTGAAGCCCAAG  
ATGATCGCGGCATCGCGGCTTATCAAGGTGGCCAGTACGACACAGATCCTGATCGAGATCTGGGCAAGAGGCCATCGGCACCGTCTGTGGGCC  
CACCCTCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGCTGCACTTCAACTTCCCCATCTCCCCATCGAGACCGTGCCTGACCTGAAGC  
CCGGCATGGACGGCCCCAAGGTGAAGCATGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGCAAGGAGATGGAGAAGGAGGCAAGATC  
TCCAAGATCGCCCCGAGAACCTTACAACACCCCTGTTCCCATCAAGAAGAGACTCCACCAAGTGGCGCAAGTGGTGGACTTCCGCGAGCTGAA  
CAAGCGACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCCCACCCCGGCCCTGAAGAAGAAGTCCGTGACCGTCTGGACGTGGCGGACCGCT  
ACTTCTCCGTGCCCTGGACCAAGACTTCCGCAAGTACACCGCTTCAACCTCCCTCCACCAACACGAGACCCCGCATCCGCTACCACTACAACG  
CTGCCCCAGGGCTGAAGGGTCCCCCGCATCTTCCAGTCTTCCATGACCAAGATCTTGAGCGCTTCCGAGACCAACCCCGAGATCTGTATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGGAGCACCTGCTGCGCTGGGGCTTACCA  
CCCCGACAAGAAGCACCAAGAGGAGCCCCCTTCTGTGGATGGCTACGAGCTGCACCCGACAAGTGGACCGTGCAGCCCCATCGTGTGCCCGAGAAG  
GACTCTGGACCGTGAACGACATCCAGAAGCTGTTGGCAAGCTGAATGGGCTCCAGATCTACGCCGGCATCAAGGTGGCCAGCTGTGCAAGTGTCT  
GGCGGGCGCAAGGCCCTGACCGAGGTATCCCCCTGACCGCGAGCGGAGCTGGAGCTGGCGAGAACCGGAGATCTTGAAGGAGCCCCGTGCACGGCG  
TGTACTAGACCCCTCCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGCCAGGCCAGTGGACCTACCAGATCTACAGGAGCCCTTCAAGAACCTGAAG  
ACCGCAAGTACGCCCGCTGCGGGCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGTCCATCGTGATCTGGGG  
CAAGACCCCAAGTTCAAGTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCGCTTGGACGCCACCTGGATCCCCGAGTGGGAGTTCGTGA  
ACACCCCCCTCTGGTGAAGCTGTGTACCAGCTGGAGAAGGAGCCCATCGTGGCGCCGAGACCTTCTAGTGGACGGCGCGGCCAACCGCGAGACCAAG  
TCCGGCAAGGCCGGTACGTGACCGACCGCGGCCGCCAGAAGTGTCTCCCTGACCGACACCAACCAAGAACCGAGCTGCAGGCCATCCACCTGGC  
CCTGCAGGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCAGTAGCCCTGGGCATCATCCAGSCCAGCCGACAACTCCGAGTCCGAGCTGGTGT  
CCCAGATCATCGACGACTGATCAAGAAGGAGAAGTGTACCTGGCTGGGTGCCCGCCACAAAGGCATCGCGCGCAACGAGCAGGTGGACAAGTGGTG  
TCCGCGGCAATCCGAAGTGTGTTCTTGACGGCATCGACAAGGCCAGGAGGCCACAGAGAATCACTCCAATGGCGGCCCATGGCTCCGACTT  
CAACTGCCCCCGTGGTGCCAAAGGATCGTGGCTCTCTGGACAAGTGCACGTGGCCTCCAGCTGAAGGGCGAGGCCATGACGGCCAGGTGACTGCTCCCCGGCA  
TCTGGCAGCTGACTGCACCCACTGGAGGGCAAGATCATCTGTGGCTGACCTGGCCCTGAAAGCTGGCCGCGCTGGCCCTCAACTTCATCTCCACCGCCGTGAA  
GGCAGGAGACCGCTACTTCTGCTGAAGCTGGCCGAGGTTCGGCATCCCCTACAACCCCTGTAAGATCATCCACCCGACAGTCCATGAACAAGCAGCTGAAGCAGA  
GGCGCTCTGTGGTGGCGGCATCAAGCAGGAGTTCGGCATCCCCTACAACCCCTGTAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGCAGCTGAAGCAGA  
TCTATCGGCCAGTTCGGGACCGCCGAGCACCTGAAGACCGCGCTGCAGATGGCGTGTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGGCGGCTAC  
TCCGCGCGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCCAGAACTTCCGCGTGTACTACCG  
CGACTCCCGGACCCCATCTGGAAGGGCCCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCTGTGTATCCAGGACAACAACGACATCAAGGTGGTGGCCCC  
GCCGAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGCGACGACTGCGTGGCTTCCCGCAGGACGAGGACTAA

Fig. 121A

## 78. 2003\_CON\_04\_CPX\_pol.PEP

FFRENVAFQQREARKESSEQARANSPPARRELDERGDNLLSEAGTEGQGTISFNFPQITLWQRPPLVTIKIGGQIREALLDTGADDTVLEEN  
 LPGKWKPKMIGGIGGFIVKVRQYDQIPIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKNSTRWRKLVDFRELNKRQDFWEVQLGIPHPAGLKKKKSVTVLVDGDAYFSVPLDP  
 EFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSIPAIFQCSMTKILEPFRTKNPEIYIYQYMDLGVGSDLEIGQHRAKIEELREHLLRWGF  
 STPDKKHQKEPPFLWMGYELHPDKWTVQPIQLAEKDSWTVDIQLVGLKNWASQIYPGKVKQLCKLLRGAKALTDIVPLTTEAELELAEN  
 REILKEPVHGAYYDPSKDLIAEIQKQGQGWYQIYQEPYKNLKTGYAKTRSAHTNDVRQLTEAVQKIAMECIVIWGKTPKFRLP IQKETW  
 DTWTEYWQATWIPWEFEVNTPLVLKLYQLETDPIAGAEFYVDGAASRETKQKAGYVTDGRQKVVSLSSETTNQKTELQAIYLALQDSG  
 SEVNIIVTDSQYALGIIQAQPDSESDLVNQIIIEQLIQKDKVYLSWVPAHKGIGGNEQVDKLVNSGIRKVLFLDGIDKAQEEHEKYHNNWRAM  
 ASDFNLPVVAKEIVASCNKCQLKGEAMHGQVDCSPGIWQLDCTHLEGIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKI IH  
 TDNGSNFTSAAVKAACWWADIQEFEGIPYNPQSQGVESMNKELKKIIGQVRDQAEHLKTAVMQMAVFIHNEFKRKGIGGYSAGERIIDIIAS  
 DIQTKELQKQITKIQNFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVP RRKAKIIRDYGKQMGAGDDCVAGRQDED\$

Fig. 122A

## 79. 2003\_CON\_06\_CPX\_pol.PEP

FFRENLAFOQGEAREFSSEQARANSPTRRRELVRRGDSPLPEAGAEQGGAISLSFPQITLWQRPPLVTVRIGGQIEALLDTGADDTVLEDIN  
 LPGKWKPKMIGGIGGFIVKVRQYDQIPIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKNSTRWRKLVDFRELNKRQDFWEVQLGIPHPAGLKKKKSVTVLVDGDAYFSVPLDE  
 DFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSIPAIFQCSMTKILEPFRTKNPEIYIYQYMDLGVGSDLEIGQHRAKIEELREHLLRWGF  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKDSWTVDIQLVGLKNWASQIYPGKVKQLCKLLRGAKALTDIVPLTAEAELELAEN  
 REILKEPVHGAYYDPSKDLIAEIQKQGQGWYQIYQEPHKNLKTGYARIKSAHTNDVKQLTEAVQKIALESIVIWGKTPKFRLP IQKETW  
 ETWTEYWQATWIPWEFEVNTPLVLKLYQLETEPIVGAETFYVDGAANRETKKAGYVTDGRQKVVSLSSETTNQKTELQAINLALQDSG  
 SEVNIIVTDSQYALGIIQAQPDKSESELVNQIIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGIDKAQEDHERYHSNWRAM  
 ASDFNLPVVAKEIVASCNKCQLKGEAMHGQVDCSPGIWQLDCTHLEGIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH  
 TDNGSNFTSAAVKAACWWANITQEFEGIPYNPQSQGVESMNKELKKIIGQVRDQAEHLKTAVMQMAVFIHNEFKRKGIGGYSAGERIIDIIAS  
 DIQTKELQKQITKIQNFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSEIKVVP RRKAKIIRDYGKQMGAGDDCVAGRQDED\$





**Fig. 122B**

2003 CON 06 CPX pol.OPT

TTCCTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGAGTTCTCTCCGAGCAGGCCCGCGCAACTCCCCCACTCCCGCCGCGAGCTGCGCGTGCG  
CCGCGGACATCCCCCTGCCGAGGCCGCGCCGAGGGCCAGGCATCTCCTGTCTTCCCAGATCACCTGTGGCAGCGCCCCCTGGTGACCG  
TGCGATCGCGCGCAGCTGATCGAGGCCCTGCTGGACACCGCGCGGACGACCCTGTGGAGACATCAACCTGCCCGCAAAGTGAAGCCCCAAGATG  
ATCGCGGCATCGCGCGCTTCATCAAGGTGCGCCAGTACGACCATCTGTATCGAGATCTGCGCAAGAAGGCCATCGGCACCGTGTGTGGGCCCCAC  
CÓCGTGAACATCATCGCGCGCAACATGCTGACCCAGATCGGCTGCACCTGAACCTTCCCATCTCCCATCGAGACCGTGTCCCGTGAAGCGCCG  
GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCGTGACCGAGGAGAAAGATCAAGGCCCTGACCGGAGATCGACCGGATGGAGAGGGAAGATCTCC  
AAGATCGGCCCGGAGAACCCCTACAACACCCCCCATCTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGAAGCTGTGGACTTCCGCGAGCTGAACAA  
GGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCGCTGAAGAAGAAGTCCGTGACCCGTCTGGAGCTGGCGAGCGCCCTACT  
TCTCCGTGCCCTGGACGAGGACTTCGCAAGTACACCGCTTCACCATCCCCCTCCATCAACAACGAGACCCCGGCATCCGCTACAGTACAACGTGCTG  
CCCCAGSCTGGAAGGGTCCCCCGCATCTTCCAGTCTCCATGATCAAGATCTTGGAGCCCTTCCGCATCAAGAACCCCGAGATCGTGATCTACCAAGTA  
CATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGGAGACCTGCTGAAGTGGGGCTTTCACCAACC  
CCGACAAAGAAGACACAGAAAGGAGCCCCCTTCTGTGGATGGGCTACGAGCTGCACCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGACAAGGAC  
TCTTGGACCGTGAACGACATCCAGAAAGCTGGTGGCAAGCTGAACCTGGGCTCCAGATCTACCCCGCATCAAGTGAAGCAGCTGTGCAAGCTGCTGCG  
CGCGCCAAAGGCCCTGACCGACATCGTGGCCCTGACCGCGGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGTGCACGGCGTGT  
ACTACGACCCCTCCAAGGACCTGATCGCGGAGATCCAGAAAGCAGGGCCAGGCGGAGTGGACCTACCAGATCTACGAGAGCCCAACAAGAACCTGAAGACC  
GGCAAGTAGCCCCGCATCAAGTCGGCCACACCAACGACGCTGAAGCAGCTGACCGAGGCCGTGCAGAAAGATCGCCCTGGAGTCCATCGTGATCTGGGGCAA  
GACCCCAAGTTCGGCTGCCCATCCAGAAAGGAGACCTGGGAGACCTGGTGAGCCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACA  
CCCCCCCCCTGGTGAAGCTGTGTACAGCTGGAGACCGAGCCCATCGTGGGCGCGAGACCTTCTACGTGACGCGCGCCCAACCGCGGAGACCAAGAAG  
GGCAAGGCGCGCTACGTGACCGACCGCGCGCGCGCAGAAAGTGTGTCTTACCGGAGACCAACCAAGACCGAGCTGCAGGCCATCAACCTGGCCCT  
GCAGGACTCCGGCTCCGAGTGAACATCGTGACCGACTCCAGTACGCGCTGGGCATCATCCAGGCCAGCCGACAGTCCGAGTCCGAGTGGTGAACC  
AGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGTCTTGGTGGCGCGCCCAACAAGGCCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTGTC  
ACCGGCATCCGCAAGGTGCTGTCTTGGACGGCATCGACAAGGCCACGAGGACCAACGAGCGCTACCACTTCAAACCTGGCGCGCATGGCCTCCGACTTCAA  
CCTTGCCCCCATCTGTGGCCAAAGGAGATCGTGGCTCTTCGCAACAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCT  
GGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGTGGCGTGCAGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGGAGACCGGC  
CAGGAGACCGCTACTTCTCTGAAGCTGGCCGCGCGTGGCCGTGAAGTGTATCCACACCGACAACGGCTCCAACCTTCACTTCCGCCCGCTGAAGGC  
CGCTGTGTGGGCCAACATCACCCAGGAGTTCGGCATCCCCCTACAACCCCGAGTCCOAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
TCGGCCAGGTGCGCGACCAAGCCGAGACCTGAAGACCGCGCTGCAGATGGCCGTGTTCATCCCAACTTCAAGCGCAAGGGCGGCATCGCGGCTACTCC  
GCCGGGAGCGCATCATCGACATCATCGCTCCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCCAGAACTTCCGCGTGTACTACCGGA  
CTCCCGGACCCCATCTGGAAGGCCCGGCGGCTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGGCTGGCCCCG  
CCAAAGGCCAAGATCATCCGGACTACGGCAAGCAGATGGCCGGCAGCATGGTGGCGCGCCGACGACGAGGACTAA

171/178

Fig. 123A

80. 2003 CON 08 BC pol. PEP  
 FFREILAFQGEAREFPPEQTRANSPTSRELQVRGDNPSSEAGTERQGTNLNFPQITLWQRPLVSIKVGGQIKEALLDTGADDTVLEEVNLP  
 KWPKMIGGIGGFIKVRQYEQPIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPIETVPVKLPGMDGPKVKQWPLTEEEKIKA  
 LTAICDEMEKEGKITIKIGPDNPYNTPIFAIRKDDSSKWRKLVDFRELNKRQDFWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVPLDKDFR  
 KYTAFTIPSVNNETPGIRYQYNVLPQGWKGSPIAFCQSMTKILEPFRKQNPDIYIYQYMDLLYVGSDELIGQHRKIEELREHLLKWGETTP  
 DKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGLNWSAQIYPGKVRQLCKLLRGAKALTDIVPLTEEALELEAENREI  
 LKEPVHGAYYDPSKELIAEIQKQGDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKIPKFRPLPIQKETWETW  
 WTDYWQATWPEWEFVNTPPLVKLWYQLEKDPAGVETFYVDGAANRETKIGKAGYVTDGRKKIVSLDTTNNQKTELQAIYIALQDSGSEV  
 NIVTDSQYALGIIQAQPDKSESELVNQII EQLIKKERVYLSWPAHKGIGGNEQVDKLVNSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASD  
 FNLPIVAKELIVASCDQCQLKGEAMHGQVDCSPGIWQLDCTHLEGGIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHTDN  
 GSNFTSAAVKAACWAGIQQEFGIPYNPQSQGVVESMNKELKKLIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIVDIIATDIQ  
 TRELOKQIIKIQNFVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIKDYGKQMGADCVAGRQDED\$

Fig. 124A

81. 2003 CON 10 CD pol. PEP  
 FFRENLAFFQQRKARELPSEQTRANSPTSRELRVWGGDNTLSETGAERQGAVALSFPQITLWQRPLVTVKIGGQKKEALLDTGADDTVLEEMN  
 LPGWKPKMIGGIGGFIKVRQYDQILIEICGYKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPIETVPVKLPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISRIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRQDFWEVQLGIPHPAGLKKKSVTVLVDGDAYFSVPLYE  
 DFRKYTAFTIPINNTPGIRYQYNVLPQGWKGSPIAFCQSMTKILEPFRKQNPDIYIYQYMDLLYVGSDELIGQHRKIEELRGHLLKWGF  
 TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGLNWSAQIYPGKVRQLCKLLRGAKALTDIVPLTEEALELEAEN  
 REILKEPVHGVYDPSKDLIAEIQKQGDQWTYQIYQEPHKNLKTGKYAKRRTAHTNDVKQLTEAVQKIAQESIIVWGTTPKFRPLPIQKETW  
 ETWWTDYWQATWPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKIGKAGYVTDGRQKVISITDTTNQKTELQAINLALQDSG  
 SEVNIIVTDSQYALGIIQAQPDKSESELVNQII EQLIKKEKYLWSVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHNNWRAM  
 ASDFNLPVPVAKELIVASCDKCQLKGEALHGQVDCSPGIWQLDCTHLEGGVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVH  
 TDNGSNFTSAAVKAACWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFRRKGGIGGYSAGERIIDIAT  
 DIQTKELQKQIIKIQNFVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIKDYGKQMGADCVASRQDEDQ



Fig. 124B

2003\_CON\_10\_CD pol.1.OPT

TTCTCCGCGAGAACCTGGCCTTCCAGCAGCGCAAGCCCGGAGCTGCCCTCCGAGCAGACCCGGCCCAACTCCCCACCTCCCGGAGCTGCGCGTGTG  
 GGGCGGCGACAACACCTGTCCGAGACCGGCGGAGCGCCAGGGCGCGTGTCCCTGTCTTCCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGACCG  
 TGAAGATCGGCGGCGCAGCTGAAGGAGGCCCTGCTGGACACCGCGCGCCGACACACCGTGTGGAGGAGATGAACCTGCCCCGCAAGTGAAGCCCCAAGATG  
 ATCGGGCGCATCGGCGGCTTCATCAAGGTGGCCAGTACGACCCAGATCCTGATCGAGATCTGCGGCTACAAGGCCATCGGCACCGTGTGGTGGGCCCCAC  
 CCCCCTGAACATCATCGGCCGCAACCTGTGACCCAGATCGGCTGCAACCTGAACCTTCCCCATCTCCCCATCGAGACCGTGTGAACTGAAGCCCCG  
 GCATGACGGCCCCAAGTGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGGATCTGCACCGAGATCTGGTGGACTTCCGCGAGCTGAACAA  
 CGCATCGGCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAAGAAGAGGACTCCACCAAGTGGCGCAAGTCCGTGCTGGACGTGGCGGACGCCCTACT  
 GCGACCCAGGACTTCTGGAGGTGCAGTGGGCATCCCCACCCCGCGGCTGAAGAAGAAAGTCCGTGACCGTCCGTGCTGGACGTGGCGGACGCCCTACT  
 TCTCCGTGCCCTGTACGAGGACTTCCGCAAGTACACCGCCTTCAACATCCCCCTCCATCAACACGAGACCCCCCGCATCCGTACCAACCGTGTG  
 CCCCAGGCTGGAAGGGTCCCCCGCATCTTCCAGTCTCCATGACCAAGATCTTGAGCCCTTCCGCAAGCAGAACCCCCGAGATGGTGTATCTACCAAGTA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCATCAAGATCGAGGAGTGGCGGCCACCTGTGAAGTGGGCTTCAACACCC  
 CCGACAAGAGCACCAAGAGGCCCTTCTGTGGATGGGTACGAGTGCACCCGACAAGTGGACCGTGCAGCCCATCCAGTGGCCCGAGAGGAC  
 TCCTGGACCGTGAACGACATCCAGAAAGCTGGTGGGCAAGTGAACCTGGCCTCCAGATCTACCCCGCATCAAGGTGGCGCAGCTGTGCAAGCTGTGCG  
 CGGCGCAAGGCCCTGACCGACATCGTGCCTGACCCGAGGAGCGGAGCTGGAGTGGCCGAGAACCGCGGAGATCTGAAGGAGCCCGTGCACGCGCTGT  
 ACTACGACCCCTCAAGGACCTGATCGCCGAGATCCAGAAGCAGGCGCAGGACCACTGGACCTACCAAGATCTACAGGAGCCCCACAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGCGCGCACCGCCACACCAACGACGTGAAGCAGTGAACGAGCTGACCGAGGCCGTGAGAAGATCGCCAGGAGTCCATCTGTGATCTGGGGCAA  
 GACCCCAAGTTCGCGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCCTGGTGAAGCTGTGTACCAAGTGGAGAGGAGCCCATCTGTGGCGCGCGAGACCTTCTAGTGGACGGCGGCCCAACCGCGAGACCAAGCTG  
 GGCAAGGCGGCTACGTACCGACCGCGCGCGCGCAGAGGTGATCTCATACCGACACCAACCAAGACCGAGCTGCAGGCCATCAACCTGGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACCAAGTCCGAGTCCGAGCTGGTGAACC  
 AGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGTCTGGTGGCGCCAGGAGGACGAGAGTACCAACAACACTGGCGGCCATGGCTCCGACTTCAA  
 TCCGGCATCCGCAAGGTGCTGTCTTGACGGCATCGACAAGGCCAGGAGGACGAGAGTACCAACAACACTGGCGGCCATGGCTCCGACTTCAA  
 CCTGCCCCCGTGGTGGCAAGGAGATCGTGGCTCTCTGGACAAGTGCAGTGAAGGGCGAGGCCCTGCACGGCCAGGTGGACTGTCCCCCGGCATCT  
 GGCAGCTGGACTGCACCACTGGAGGGCAAGGTGATCTGTGGTGGCGTGCAGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCGCGGAGACCGGC  
 CAGGAGACCGCTACTTCTGCTGAAGCTGGCCGCGCGTGGCCGTGAAGGTGTGCACACCGACCAACGGCTCCAACCTCACCTCCGCGCGCGTGAAGGC  
 CGCCTGTGTGGCCGGCATCAAGCAGGAGTTCGGCATCCCCACAACCCCAAGTCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGCGGACCAAGCCGAGCACTGAAGACCGCGCTGCAGATGGCCGTGTTCATCCACAACCTCAAGCGCAAGGGCGGCATCGGCGGCTACTCC  
 GCGCGGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCCAGAACCTCCGCGTGTACTACCGCGA  
 CTCCCGGACCCCATCTGGAAGGGCCCCCGCAAGCTGTGTGAAGGGCGAGGGCGCGTGGTGTATCCAGGACAACCTCCGACATCAAGGTGGTGGCCCCGCC  
 GCAAGGTGAAGATCATCAAGGACTACGGCAAGCAGATGGCCGCGCGGACTGCGTGGCTTCCCCGCCAGGACGAGGACCCAG

173/178

Fig. 125A

## 82. 2003 CON 11 CPX pol. PEP

FFRENLA~~F~~QGE~~A~~RE~~F~~SPEQARANSPTSRELVRGGDSPLPETGAEGE~~G~~AISFNFPQITLWQRP~~L~~VTIKVAGQLKEALLDTGADDTVLEED  
 L~~P~~GRWKPKMIGGIGGFIKVRQYEEIIIEIEGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIDTVPVKLKPGMDGPKVKQWPLTEEK  
 IKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKR~~T~~QDFWEVQLGIPHPAGLKKKSVTVL~~D~~VGDAYFSVPLDE  
 SFRKYTAFTIP~~S~~INNETPGIRYQYNVLPQGWKGS~~P~~AI~~F~~QSSMTKILEPFR~~T~~QNP~~E~~I~~V~~IYQYMD~~D~~LYVGS~~D~~LEIGQHREKVEELRKHL~~L~~KKWGF  
 TTPDKKHQKEPPFLMWGYELHPDKWTVQPIQLPDKECWTVNDIQLVGKLNWASQIYPGIKVKQ~~C~~KLRLRGTKALTDI~~V~~PLTAAEAELELAEN  
 REILKEPVHGVYDPSKDLIAEVQKQGLDQWTYQIYQEPFKNLKTGYAKRR~~T~~AHTNDVRQLAEVVQKISMESIVIWGKIPKFR~~L~~PIQRETW  
 ETWTDYQATWIP~~E~~WEFVNT~~P~~PLVLWYQLEKEPIIGAETFYVDGAANRET~~K~~LKGAGYVTDKGRQKVTVLTETTNQKTELEA~~I~~H~~L~~ALQDSG  
 LEVNI~~V~~TD~~S~~QYALGIIQAOPDKSESELVSQIIELIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEH~~E~~RYHNSNWRAM  
 ASDENLPPIVAK~~E~~IVASCDKQ~~L~~KGEAMHGQVDCSPGIWQ~~L~~DC~~T~~HLEGIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH  
 TDNGSNFTSA~~A~~VKAACW~~W~~ANIQQEFGIPYNPQSQGVVESMNKELKKIIQVREQA~~E~~HLKTAVQMAVFIH~~N~~FRKGGIGGYSAGERIVDIIAT  
 DLQTKELQKQITKIQNFRVYRDSRDP~~I~~WKGPAKLLWKGE~~G~~AVVIQD~~N~~SDIKV~~V~~PRRKA~~K~~IIIRDY~~G~~KQ~~M~~AGDDC~~V~~AGRQDE~~D~~\$

174/178

Fig. 126A

## 83. 2003 CON 12 BF pol. PEP

FFRENLA~~F~~QGE~~A~~RK~~F~~PSEQARANS~~P~~ASRELWVR~~R~~GD~~N~~PLSEAGAERRGTVP~~S~~LSFPQITLWQRP~~L~~VTIKVGGQLKEALLDTGADDTVLEDI  
 NLPGKWKPKMIGGIGGFIKVKQYDNILIEICGHKAIGTVLVGPTPVNIIGRNLLTQLGCTLNFPISP~~I~~ETVPVKLKPGMDGPKVKQWPLTEE  
 KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRKLVDFRELNKR~~T~~QDFWEVQLGIPHPAGLKKKSVTVL~~D~~VGDAYFSVPLD  
 KDFRKYTAFTIPSVNNETPGIRYQYNVLPQGWKGS~~P~~AI~~F~~QSSMTKILEPFRKQNPDI~~V~~IYQYMD~~D~~LYVGS~~D~~LEIGQHRTKIEELRQHLLRWG  
 FTT~~P~~DKKHQKEPPFLMWGYELHPDKWTVQPIVLPEKDSWTVNDIQLVGKLNWASQIYPGIKVKQ~~C~~RLRLRGTKALTEVI~~P~~LTKEAELELAE  
 NREILKEPVHGVYDPSKDLIAEIQKQGGQWTYQIYQEPFKNLKTGYARMGAHTNDVKQ~~L~~TEAVQKITTESIVIWGKTPKFR~~L~~PI~~L~~KET  
 WDTWTEY~~W~~QATWIP~~E~~WEFVNT~~P~~PLVLWYQLETEPIAGAETFYVDGASNRET~~K~~KGAGYVTDGRQKAVSLTETTNQKAE~~L~~HAIQALQDS  
 GSEVNI~~V~~TD~~S~~QYALGIIQAOPDKSESELVNQIIELIKKEKVYLSWVPAHKGIGGNEQVDKLV~~S~~AGIRKILFLDGIDKAQEEH~~E~~KYHNNWRA  
 MASDENLPPVAK~~E~~IVASCDKQ~~L~~KGEAMHGQVDCSPGIWQ~~L~~DC~~T~~HLEGIILVAVHVASGYLEAEVIPAETGQETAYFILKLAGRWPVKTI  
 HTDNGPNFSSAAVKAACW~~W~~AGIQEFGIPYNPQSQGVVESMNKELKKIIQV~~R~~DQA~~E~~HLKTAVQMAVFIH~~N~~FRKGGIGGYSAGERIIDIIIS  
 TDIQTRELQKQIIKIQNFRVYRDSRDPVWKGPAKLLWKGE~~G~~AVVIQD~~N~~SEIKV~~V~~PRRKA~~K~~IIIRDY~~G~~KQ~~M~~AGDDC~~V~~AGRQDE~~D~~\$



Fig. 125B

2003\_CON\_11\_cpx\_pol.OPT

TTCTTCGGCAGAACCTGGCCCTCCAGAGGGCGAGGCCCGGAGTTCTCCCCGAGCAGGCCCGGCCAACTCCCCACCTCCCGCGAGCTGCGCGTGCG  
 CGCGGCGACTCCCCCTGCCGAGACCGGCGGAGGGCGGCCATCTCTTCAACTTCCCCAGATCACCTGTGGCAGCGCCCCCTTGGTGACCA  
 TCAAGGTGGCGGCCAGCTGAAGAGGCCCTGTGGACACCGGCGCCGACACCGTGTGGAGGAGATCGACCTGCCCCGGCCGCTGGAAGCCCAAGATG  
 ATCGGCGGCATCGCGGCTTCATCAAGGTGCGCCAGTACGAGGAGATCATCGAGATCGAGGGCAAGAGGCCATCGGCACCGTGTGGTGGGCCCCAC  
 CCCCCTGAACATCATCGGCGCAACATCTGACCCAGATCGGCTGCACCTGAACCTTCCCCATTCGCCATCGACACCGTGTGGTGAAGCCCCG  
 GCATGGACGGCCCCAAGTGAAGCAGTGGCCCCGTGACCGAGGAGAGATCAAGGCCCTGACCGAGATCTGACCGAGATGGAGAGGGCAAGATCTCCGAGCTGAACAA  
 AAGATCGGCCCCGAGAACCCCTAACACACCCCGTGTTCGCCATCAAGAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAGCTGAACAA  
 GCGCACCCAGGACTTCTGGAGGTGAGCTGGGCAATCCCCACCCCGCGCCCTGAAGAGAAAGTCCGTGACCGTGTGGACGTGGCGACGCTACT  
 TCTCCGTGCCCTGGACGAGTCTTCGCAAGTACACCGCTTCAACCTCCCATCCCCCTCCATCAACACGAGACCCCCGGCATCCGCTACCACTAACACGTG  
 CCCCAGGCTGGAAGGCTCCCCGCCATCTCCAGTCTCCATGACCAAGATCTTGAGCCCTTCCGACCCCAAGAACCCCGAGATCGTGATCTACCACTA  
 CATGGACGACCTGTACGTGGGTCCGACCTGGAGATCGGCGAGCACCCGAGAGAGTGGAGGAGCTGGCAAGCACCTGCTGAAGTGGGCTTCACCAACCC  
 CCGACAAGACCAAGAGAGCCCCCTTCTGTGGATGGGTACGAGCTGCACCCGACAAAGTGGACCGTGCAGCCCCATCCAGCTGCCCGACAAGGAG  
 TGCTGGACCGTGAACGACATCCAGAGCTGGTGGCAAGCTGAACCTGGGCTCCAGATCTACCCGGCATCAAGTGAAGTGAAGCAGCTGTGCAAGCTGTGG  
 CGGCACCAAGGCCCTGACCGACATCGTGCCCTGACCGCGAGGCCGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTTGAAGGAGCCCCGTGACCGCGTGT  
 ACTACGACCCCTCCAGGACCTGATCGCCGAGTGCAGAGCAGGCGCTGGACCACTGACCACTACCAAGTGGTGCAGAGATCTCCATGGAGTCCATCGTGATCTGGGGCAA  
 GGCAAGTACGCCAAGCGCCGACCGCCACACCAAGACCTGGCCAGCTGGCCGAGGTGGTGCAGAGATCTCCATGGAGTCCATCGTGATCTGGGGCAA  
 GATCCCAAGTTCGGCTGCCATCCAGCGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACA  
 CCCCCCCTGGTGAAGCTGTGGTACCACTCCAGCGGAGAGCCCATCATCGGCGCGAGACCTTCTACGTGGACGGCGCCGCAACCGGAGACCAAGCTG  
 GGCAAGGCCGGTACGTGACCGACAAGGGCGCCAGAGGTGTGACCCCTGACCGAGACCAACCAAGAGACCGAGCTGGAGGCCATCCACCTGGCCCT  
 GCAGGACTCCGGCTGGAGTGAACATCGTGACCGACTCCAGTACGAGTGGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGTGTCCC  
 AGATCATCGAGCAGTGCATCAAGAGGAGAGGTGTACCTGTCTGGTCCCCGCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAAGTGGTGTCC  
 TCCGGCATCCGCAAGGTGTCTTCTGGACGGATCGACAAGGCCAGGAGGACGAGCGCTACCACTCCAACTGGCGGCCCATGGCCTCCGACTTCAA  
 CCTGCCCCCATCGTGGCCAAAGAGATCGTGGCTCTTCGACAAAGTGCAGCTGAAGGGCAGGCCATGCAGGCCAGGTGGACTGCTCCCCCGGCATCT  
 GGAGCTGGACTGCACCCACTGGAGGGCAAGATCATCTGTGGCCGTGCACGTGGCCCTCCGGTACATCGAGGCCGAGGTGATCCCCGCCGAGACCGGC  
 CAGGAGACCGCCTACTTCTGAGCTGGCCCGCGTGGCCCGTGAAGGTGATCCACACGACACCGGCTCCAACCTTACCTCCGCGCCGTGAAGGC  
 CGCCTGTGTGGGCCAACATCCAGCAGGATTCGGCATCCCCCTAACCCCCAGTCCAGGGCTGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCCAGGTGGCGGAGCAGGCCGAGCACCTGAAGACCGCGTGCAGATGGCCGTGTTCATCCACAATTCGAAGCCGAGGGCGCATCGGGGGTACTCC  
 GCCGGCAGCGCATCGTGGACATCATCGCCACCGACCTGCAGACCAAGGAGCTGCAGAAAGCAGATCACCAAGATCCAGAACTTCCGCGTGTACTACCGCGA  
 CTCCCCGACCCCATCTGGAAGGGCCCCGCCAAGCTGTGTGGAAAGGGCGGCCGTGGTGTATCCAGACAACTCCGACATCAAGGTGGTGGCCCCGCC  
 GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGCGGACGACTGCTGTGGCCGCCCGCCAGGACGAGGACTAA

175/178

176/178

## Fig. 126B

2003\_con\_12\_BF\_pol.1.OPT

TTCTTCGGGAGAACTGGCCCTTCAGAGGGGAGGCCCGCAAGTTCCCTCCGAGCAGGCCCGCGCCAACTCCCCCGCTCCCGCGAGCTGTGGGTGCG  
CCGGCGGACAAACCCCTGTCCGAGGCGGCGGAGCGCGGCAACCGTGCTCCCTTCCCCAGATCACCTGTGGCAGCGCCCTTGGTGA  
CCATCAAGGTGGCGGCCAGCTGAAGGAGGCCCTGCTGGACACCGGGCGCGACACCGTGCTGGAGGACATCAACCTGCCCGCAAGTGAAGCCCAAG  
ATGATCGCGGCATCGCGGCTTCATCAAGGTGAAGCAGTACGACAACATCCTGATCGAGATCTGGGCCACAAGGCCATCGGCACCGTGCTGGTGGGCC  
CACCCCGTGAACATCATCGGCCGCAACCTGCTGACCAAGTGGCTGCACCTGAACCTTCCCATTCGCCCATCGAGACCGTGCCCGTGAAGCTGAAGC  
CCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGACCGAGATGGAGAAGGAGGCAAGATC  
TCCAAGATCGGCCCGGAGAACCCCTACAACACCCCGCTGTTCCGCATCAAGAAGAGGACTCCACCAAGTGGCGCAAGCTGGTGGAATTCGCGAGCTGAA  
CAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCACCCCGCGGCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGCGGACGCTT  
ACTTCTCCGTGCCCTGGACAAGGACTTCGCAAGTACACCGCTTACCATCCCTCCGTGAACAACGAGACCCCGCATCCGCTACCGTATCAACATG  
CTGCCCCAGGGCTGGAAGGCTCCCCCGCATCTTCCAGTCTCCATGACCAAGATCCTGGAGCCCTTCGCAAGCAGAACCCCGACATCGTGATCTACCA  
GTACATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGCAACCAAGATCGAGGAGTGGCCAGCACCTGCTGGCTGGGCTTCACCA  
CCCCGACAAGAAGCACCAAGAGGAGCCCCCTTCTGTGGATGGCTACGAGTGCACCCCGACAAGTGGACCGTCAAGCTCGTGTGCCCGAGAAAG  
GACTCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCTCCAGATCTACCCCGCATCAAGTGAAGCTGACCGCTGTGCCCGCTGCT  
GCGCGCACCAAGGCCCTGACCGAGGTGATCCCCCTGACCAAGGAGGCCGAGCTGGAGTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGTGCAAGGCG  
TGTAATACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGCCAGTGGAATCCAGATCTACAGGAGCCCTTCAAGAACCTGAAG  
ACCGGCAAGTACGCCCGCATGCGCGGCGCCACACCAAGCAGTGAAGCAGTGAACCGAGGCGCTGCAGAAATCACCAACCGAGTCCATCGTGATCTGGG  
CAAGACCCCCAAGTTCGGCTGCCATCTGAAGGAGACTGGGACACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGA  
ACACCCCCCTGGTGAAGCTGTGGTACCAAGTGGAGACCGAGCCCATCGCCGCGCGCGAGACCTTCTACGTGGACGGCGCTTCAACCGCGAGACCAAG  
AAGGCAAGCGCGCTACGTGACCGACCGCGCGCGCGAGGCCGTGTCCTGACCGAGACCAACCAAGAGCGCGAGCTGCACGCCATCCAGCTGGC  
CCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGTCCGAGTCCGAGCTGGTGA  
ACCAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGATCTGTCTGGTGCCGCCCAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTG  
TCCGCCGGCATCCGCAAGATCCTGTTCCTGGACGGCATCGACAAGGCCCGAGGAGCACGAGAAGTACCAACAACCTGGCGGCCATGGCTCCGACTT  
CAACCTGCCCCCGTGGTGGCCAAAGGAGATCGTGGCTCCTGCGACAAGTGGCAGTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA  
TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCTGTGTGGCCGTGACAGTGGCTCCGGTACTCTGGAGGCCGAGGTGATCCCCGCGGAGACC  
GGCAGGAGACCGCTACTTCACTCTGAAGTGGCCGCGCTGGCCGTGAAGACCATCCACACCGACACCGGCCCACTTCTCTCCGCGCGCTGAA  
GGCCGCTGCTGGTGGCGCGCATCCAGCAGGAGTTCGGCATCCCCACAACCCCGCTGAGAGCCATCCAGGGCGTGGTGGATCCATGAACAAGGAGCTGAAGA  
TCATCCGCCAGGTGGCGGACCAAGCGGACCTGAAGACCGCGCTGAGATGGCCGTTCATCCACAACCTTCAAGCGCAAGGGCGGCATCGGCGGCTAC  
TCCGCGGGGAGGGCATCATCGACATCATCTCCACCGACATCCAGACCCGCGAGCTGCAGAAGCAGATCATCAAGATCCAGAATCTCCGCGTGTACTACCG  
CGACTCCCGGACCCCGTGTGGAAGGGCCCCGCAAGCTGCTGTGGAAGGGCGAGGCCCGCTGGTGTATCCAGGACAACCTCCGAGATCAAGGTGGTGGCCCC  
GCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCCGCGCGACGACTGCGTGGCGCGCGCGAGGAGGAGCTAA



177/178

## Fig. 127A

84. 2003 CON 14 BG pol.PEP  
FFRENLA<sup>F</sup>QOCEARE<sup>F</sup>FSPEQARANSPTRRRELWVRRGDSPLPEARAEGKGDIPLSLPQITLWQRPLVTVRIGGQLIEALLDTGADDDTVLEDIN  
LPGKWKPMIGGIGGFIVKVRQYDQILIEICGKKAIGTVLVGPTPINIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEK  
IKALTDICTEMEREGKISKIGPENPYNTPIFAIKKSDTKWRKLVDFRELNRKTQDFWEVQLGIPHPISGLKKKSVTVLVDVGDAYFSVPLDE  
SFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPIFQSSMTKILEPFRIKNPEIYIYQYMDLTVGSDLEIGQHRAKIEELRKHLLSWG  
TTPDKKHQKEPPFLMMGYELHPDKWTVQPIQLPDKESWTVNDIOKLVLGNWASQIYPGIVKQCLCKLLRGAKALTDIVPLTAEAELELAEN  
REILKEPVHGVYEP<sup>S</sup>SKELIAEVQKQGLDQWTYQIYQEPYKNLKTGKYAKRGSAHTNDVKQLTEVVQKIATESIVIWGKTPKFKLPIRKETW  
EWWTEYWOATWIPDWEFVNTPPLVKLWYRLETEPIAGAEITYYVDGAANRET<sup>K</sup>LGKAGYVTDKGKQKIIITLTETTNQKAE<sup>L</sup>QAIHIALQDSC  
SEVNI<sup>V</sup>TD<sup>S</sup>QYALGI<sup>I</sup>QAQPD<sup>R</sup>SESEV<sup>N</sup>QII<sup>E</sup>QLIKKEK<sup>V</sup>YLSWVP<sup>A</sup>HKGI<sup>G</sup>GGNEQ<sup>V</sup>DKLVSSGIRK<sup>V</sup>LFDGIDKAQEEHEKYHSNWRAM  
ASDFNLPPVVAKEIVASCDK<sup>C</sup>QLKGEAMHGQVDCSPGIWQ<sup>L</sup>DC<sup>T</sup>HLEKII<sup>L</sup>VAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKIIH  
TDNGSNFT<sup>S</sup>AAVKAAACWWANITQEFGI<sup>P</sup>YNPQSQGV<sup>V</sup>ESMNKELKKII<sup>G</sup>QVRDQAEHLKTAVQMAVFIH<sup>N</sup>FRKGGIGGYSAGERIIDIIAS  
DIQTKELQKITKIQNFRVYFRDSRDP<sup>I</sup>WKGP<sup>A</sup>KLLWKGE<sup>A</sup>VVIQDNNEIKVVP<sup>R</sup>RRKAKIIRDYGKQ<sup>M</sup>AGDDCVAGRQDED<sup>S</sup>

## Fig. 127B

2003\_con\_14\_BG\_pol.OPT

TTCTTCGGCGAGAACCTGGCCCTTCAGCAGGGCGAGGCCCCGGAGTTCTCCCCGAGCAGGCCCCGGCCCAACTCCCCACCCGCCGCGAGCTGTGGGTGCG  
 CCGGGCGACTCCCCCTGCGCGAGGCCCGCGCGAGGCAAGGGCGACATCCCTGTCTCTGCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGACCG  
 TGCGCATCGGGCGCAGCTGATCGAGGCCCTGCTGGACACCGGGCGCGACGACACCGTGTGGAGACATCAACCTGCCCGGCAAGTGAAGCCCAAGATG  
 ATCGGGCGCATCGGGCGCTTCATCAAGTGGCCAGTACGACCATCTGTATCGAGATCTGGGCAAGAGGCCATCGGCACCGTCTGGTGGGCCCCAC  
 CCCCATCAACATCATCGGGCGCAACATGCTGACCCAGATCGGCTGCACCTGAACCTCCCATCTCCCATCGAGACCGTGCCTGAAAGCTGAAGCCCG  
 GCATGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGACCGAGATGGAGCGCGAGGCAAGATCTCC  
 AAGATCGGGCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAAGAAGAAGACTCCACCAAGTGGCGAAGCTGGTGGACTTCGCGAGCTGAACAA  
 GCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCCCACCCCTCGGGCTGAAGAAGAAGTCCGTGACCGTGTGGACGTGGCGGACGCTACT  
 TCTCCGTGCCCCCTGGACGAGTCTTCCGCAAGTACCCGCTTACCATCCCTCCACCAACACGAGACCCCGGCATCCGCTACCAAGTACAACGCTGCTG  
 CCCCAGGGCTGGAAGGGCTCCCGCCATCTCCAGTCTCCATGACCAAGATCTTGAGCCCTTCGCGATCAAGAACCCCGAGATCGTGATCTACCAATA  
 CATGGACGACCTGTACGTGGCTCCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGTGCAGAACCTGCTGTCTGGGGCTTACCCACCC  
 CCGACAAGAAGCACCAAGGAGCCCCCTTCTGTGGATGGCTACGAGTGCACCCGCAAGTGGACCGTGCAGCCCATCCAGTGCCTCCGACCAAGGAG  
 TCCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCTCCAGATCTACCCGGCATCAAGTGAAGCAGCTGTCAAGCTGTGCG  
 CGCGCCAAAGCCCTGACCGACATCGTGCCTGACCGCCGAGGCGGAGCTGGAGTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGTGCACGGCGTGT  
 ACTACGAGCCCTCCAAGGAGTGTGCGCGAGGTGCAGAAAGCAGGCGCTGGACCAAGTGGACCTACAGATCTACAGAGGCCCTACAAGAACCTGAAGACC  
 GGCAAGTACGCCAAGCGGGCTCCGCCACACCAACGACGTGAAGCAGCTGACCGAGGTGGTGCAGAAGATCGCCACCGAGTCCATCGTGTCTGGGGCAA  
 GACCCCAAGTCAAGTGCCCATCCGCAAGGAGACCTGGGAGGTGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGACTGGGAGTTCGTGAACA  
 CCCCCCCTGGTGAAGCTGTGGTACCGCTGGAGACCGAGCCCATCGCCGGCGCGGAGACCTACTACGTGGACGGCGCGCCCAACCGCGAGACCAAGCTG  
 GGCAGGCCGGCTACGTGACCGACAGGCAAGCAGAAGATCATCACCTGACCGAGACCAACCAAGAGCCGAGCTGCAGGCCATCCACATCGCCCT  
 GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACCGCTCCAGTCCGAGTGGTGAACC  
 AGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGTCTGGTGGTCCCGCCCAAGGGCATCGGCGGCAACGAGAGGTGGACAAGCTGGTGTCC  
 TCCGGCATCCGCAAGGTGTCTTCTGGACGGCATCGACAAGGCCAGGAGGACACGAGAAGTACCATCCAACTGGCGGCCATGGCTCCGACTTCAA  
 CCTGCCCCCTGGTGGCAAGGAGATCGTGGCTCTTCGACAAAGTGCCAGCTGAAGGGCGAGGCCATGCAGGCCAGGTGGACTGTCCCCCGGCATCT  
 GGCAGCTGACTGCACCCACCTGGAGGGCAAGATCATCTGTGGCCGTGCACGTGGCTCCGGCTACATCGAGGCCGAGGTGATCCCCCGCGAGACCGGC  
 CAGAGACCGCCTACTTCATCTGAAGCTGGCGGGCGCTGGCCCGTGAAGATCATCCACACCGACAACGGCTCCAACTTCACTCCGCGCCGTGAAGGC  
 CGCTGTGTGGGCCAACATCAACAGGATTCGGCATCCCTTACAAACCCCAAGTCCCAAGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCA  
 TCGGCAGGTGGCGACCGAGCCGAGCACCTGAAGACCGCGCTGCAGATGGCCGTGTTCATCCACAACTTCAAGGCCAAGGCCGCGCATCGGCGGCTACTCC  
 GCGGGAGCGCATCATCGACATCATCGCTCCGCTCCGACATCCAGACCAAGAGCTGCAAGAGCAGATCAACCAAGATCCAGAAGTCCGCGTGTACTTCCGCGA  
 CTCCCGGACCCCATCTGGAAGGGCCCCCAAGCTGTGTGGAAGGGCGAGGGCGCGCTGGTGTATCCAGGACAAACACGAGATCAAGGTGTGTCGCCGCC  
 GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGCCGACGACTGCGTGGCCGGCCCGCAGGACGAGGACTAA

178/178

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**